

QCAD Simulation and Optimization of Semiconductor Quantum Dots

X. Gao,^{1, a)} E. Nielsen,¹ R. P. Muller,¹ R. W. Young,¹ A. G. Salinger,¹ N. C. Bishop,¹ M. Lilly,¹ and M. S. Carroll¹
Sandia National Laboratories, 1515 Eubank SE, Albuquerque, NM 87123, USA

(Dated: 1 April 2014)

We present the Quantum Computer Aided Design (QCAD) simulator that targets modeling multi-dimensional quantum devices, particularly silicon multi-quantum dots (QDs) developed for quantum bits (qubits). This finite-element simulator has three differentiating features: (i) its core contains nonlinear Poisson, effective mass Schrodinger, and Configuration Interaction solvers that have massively parallel capability for high simulation throughput, and can be run individually or combined self-consistently for 1D/2D/3D quantum devices; (ii) the core solvers show superior convergence even at near-zero-Kelvin temperatures, which is critical for modeling quantum computing devices; and (iii) it interfaces directly with the full-featured optimization engine Dakota. In this work, we describe the capabilities and implementation of the QCAD simulation tool, and show how it can be used to both analyze existing experimental QD devices through capacitance calculations, and aid in the design of few-electron multi-QDs. In particular, we observe that computed capacitances are in rough agreement with experiment, and that quantum confinement increases capacitance when the number of electrons is fixed in a quantum dot. Coupling of QCAD with the optimizer Dakota allows for rapid identification and improvement of device layouts that are likely to exhibit few-electron quantum dot characteristics.

I. INTRODUCTION

Silicon-based qubits for quantum information processing have attracted enormous interest and made remarkable progress^{1–4} during recent years due to long spin decoherence times and the well-established Si nanoelectronic fabrication infrastructure. There has been a wide range of research work^{5–8} in academia and national labs dedicated to design and fabricate Si-based multi-quantum dot (QD) devices for qubit application. However, many practical issues have made it challenging to achieve reliable few-electron multi-QD qubits, including fabrication process variation, interface quality, defect/disorder, and device design.

To facilitate the experimental development of the technology, a simulation tool was needed to efficiently solve for semiclassical (*i.e.*, using the Thomas-Fermi approximation) and self-consistent quantum electrostatic potentials, single- and multi-electron wave functions, and energies at near-zero temperatures. It was also essential that such a tool be able to simulate and optimize many different single- and multi-QD structures very efficiently, and provide fast feedback on which device layouts are more likely to lead to few-electron behavior. Although there have been numerous papers^{9–13} on simulation methods for quantum dots, they focus on different issues than those addressed in this work. Existing commercial¹⁴ and academic¹⁵ device simulators either target room-temperature and many-electron devices, whereas qubit applications require temperatures close to zero Kelvin and one/few-electron devices, or target simple and few geometries, whereas multi-QD devices have very complex three-dimensional (3D) shapes and can have many different layouts due to their inherently large design

space. These existing tools often have license and platform restrictions which place non-fundamental but very real limits on simulation efficiency. The Quantum Computer Aided Design (QCAD) simulator has been developed to address these challenges associated with modeling realistic multi-QDs, including complex geometries, many device layouts, low temperature operation, and 3D quantum confinement effects. It is an integrated open-source finite-element-based tool that leverages a number of Sandia-developed software programs¹⁶, including the Trilinos suite, the Albany code¹⁷, the Dakota toolbox, and the Cubit geometry and meshing tool.

The first and major part of this work details the QCAD simulator¹⁸ (or just “QCAD” for brevity), which is built upon the Albany finite element framework¹⁷ and contains three core modules of Poisson (P), Schrodinger (S), and Configuration Interaction (CI) solvers. These physical solvers can be run individually or combined self-consistently (*i.e.*, self-consistent S-P and S-P-CI solvers) for simulating arbitrary 1D/2D/3D quantum devices made from multiple different materials. They have demonstrated fast and robust convergence behavior even at very low (milli-Kelvin) temperatures. Furthermore, very high simulation throughput has been achieved using a combination of pre- and post-processing scripting, automated structure creation and meshing, and distributed parallel computing capability and resources. In Sec. II, details of the QCAD simulator and its core physical solvers are presented, with the semiclassical Poisson solver detailed in Sec. IIA, the Schrodinger solver described in Sec. IIB, and the self-consistent S-P solver detailed in Sec. IIC.

The second part of this work presents two examples of applying QCAD to simulate realistic QD structures. Since QCAD solves for the electron density at a given set of device control voltages, it is able to compute the capacitance between a quantum dot and a control gate. Such capacitances can also be determined experimentally. In

^{a)} xngao@sandia.gov

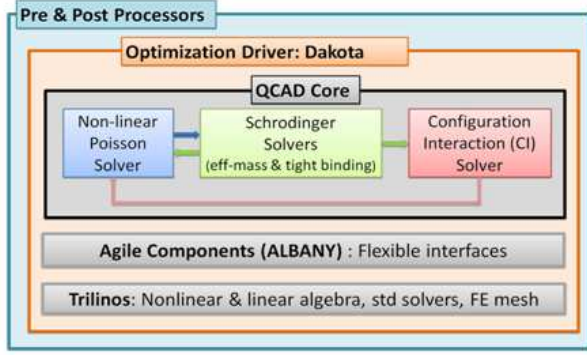


FIG. 1. (Color online) Schematic diagram of the QCAD code structure.

Sec. III, we compute the dot-to-gate capacitances in an experimental QD structure, using the semiclassical and self-consistent S-P solvers, and compare them with experiment. A detailed analytic analysis is then conducted to help understand the effect of quantum confinement on capacitances. Finally, in Sec. IV, QCAD's ability to assess and improve upon the design of nano-scale devices is demonstrated with few-electron double QDs (DQDs). Through its coupling with the optimization driver Dakota, QCAD is able to optimize gate voltages to simultaneously achieve multiple desired targets (*e.g.* few electrons in a dot and a tunnel barrier being controllable) in many different DQD device designs. We show how this aids in the identification of device designs which are likely to exhibit few-electron quantum dot behavior.

II. QCAD SIMULATOR

The QCAD code structure is given in Fig. 1. The base of the structure is Trilinos¹⁶, an open-source suite consisting of mathematical libraries (nonlinear and linear solvers, preconditioners, eigensolvers, etc.), discretization (finite element and finite volume) utilities, automatic differentiation library, distributed parallel library, and many more packages (refer to trilinos.sandia.gov for details). The Albany code¹⁷ provide a unified and flexible interface to those Trilinos packages to minimize the coding effort for users developing physical models. The Dakota driver¹⁶ interfaces with QCAD and repeatedly calls the QCAD executable to perform specified optimization tasks. It is an open-source tool that provides a broad spectrum of capabilities including nonlinear least square optimization (see dakota.sandia.gov for details). Pre- and post-processing scripts have been developed to automatically generate QCAD input files for multiple DQD devices, submit jobs to computing clusters, and collect simulation results. These scripts are one of the enabling factors that make efficient QCAD simulations possible.

The QCAD core contains nonlinear Poisson (P), effective-mass Schrodinger (S), and CI solvers. These

solvers can be run individually or combined in a self-consistent manner. The Poisson and Schrodinger solvers can be coupled self-consistently to obtain single-particle envelope wave functions and energy levels. The S-P solutions can then be used as wave function basis for the CI solver which is then coupled with the Poisson solver. This self-consistent S-P-CI technique produces accurate many-particle wave functions and energies within the effective mass approximation, which are important for quantitative study of few-electron multi-QD behavior (*e.g.* for computing the exchange energy in DQDs). Details about the Poisson, Schrodinger, and self-consistent S-P solvers are described in the following subsections, and the CI method can be found in Ref. 19. We note that the current version of the QCAD simulator solves one or a system of equations in the full number of dimensions of the problem, and does not make any separability assumptions that allow a higher dimensional problem (*e.g.* 3D) to be divided into multiple lower dimensional problems (*e.g.* 2D+1D). Such simplifying approximations²⁰ are a possible future enhancement to the QCAD core.

A. Poisson Solver

The well-known Poisson equation in a bulk semiconductor is given by

$$-\nabla \cdot (\epsilon_s \nabla \phi) = q(p - n + N_D^+ - N_A^-), \quad (1)$$

where ϕ is the electrostatic potential, ϵ_s is the static permittivity, q is the elementary charge, n and p are the electron and hole concentrations respectively, N_D^+ and N_A^- are ionized donor and acceptor concentrations respectively. Note that the form of the left hand side in Eq. (1) allows ϵ_s to have spatial dependence.

1. Carrier Statistics

n and p are given by carrier statistics for bulk (spatially unconfined) semiconductors. Both Maxwell-Boltzmann (MB) and Fermi-Dirac (FD) statistics are implemented in QCAD. For the MB statistics, n and p take the exponential forms,

$$\begin{aligned} n &= N_C \exp\left(\frac{E_F - E_C}{k_B T}\right), \\ p &= N_V \exp\left(\frac{E_V - E_F}{k_B T}\right), \end{aligned} \quad (2)$$

where k_B is the Boltzmann constant, T is the lattice temperature, E_C and E_V are the conduction and valence band edge respectively, and E_F is the extrinsic Fermi level (more details on E_C , E_V , and E_F are given in Sec. II A 2). For the FD statistics, n and p are expressed in terms of the Fermi-Dirac integrals (see Appendix A

for the derivation),

$$\begin{aligned} n &= N_C \mathcal{F}_{1/2} \left(\frac{E_F - E_C}{k_B T} \right), \\ p &= N_V \mathcal{F}_{1/2} \left(\frac{E_V - E_F}{k_B T} \right). \end{aligned} \quad (3)$$

N_C and N_V are effective density of states (DOS) in the conduction and valence band, respectively. Assuming parabolic band structure, we have

$$\begin{aligned} N_C &= 2 \left(\frac{m_n^* k_B T}{2\pi \hbar^2} \right)^{3/2}, \\ N_V &= 2 \left(\frac{m_p^* k_B T}{2\pi \hbar^2} \right)^{3/2}, \end{aligned} \quad (4)$$

where \hbar is the reduced Planck constant, m_n^* and m_p^* are respectively the electron and hole DOS effective mass including all equivalent band minima. For bulk silicon, there are six equivalent conduction minima, and the valence band minimum is degenerate including heavy hole and light hole bands at the Γ valley, hence

$$\begin{aligned} m_n^* &= 6^{2/3} (m_l m_t)^{1/3}, \\ m_p^* &= \left(m_{hh}^{3/2} + m_{lh}^{3/2} \right)^{2/3}, \end{aligned} \quad (5)$$

with m_l , m_t , m_{hh} , and m_{lh} being the electron longitudinal, electron transverse, heavy hole, and light hole effective mass, respectively.

The $\mathcal{F}_{1/2}(x)$ function in Eq. (3) is the Fermi-Dirac integral of 1/2 order and is defined as²²

$$\mathcal{F}_{\frac{1}{2}}(x) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{\varepsilon} d\varepsilon}{1 + \exp(\varepsilon - x)}. \quad (6)$$

Although the closed form of this integral can be formally expressed by the polylogarithm function²³ or by a complete expansion discussed in Ref. 24, the polylogarithm function and the complete expansion involve summations of infinite series. Hence in practice, one has to either use certain approximation to obtain a computable analytic expression, or use numerical integration techniques^{21,25}. There have been a few approximate analytic forms proposed in literature^{22,26,27} that offer relatively simple expressions and sufficient accuracy. Among them, the approximate expression in Ref. 26 takes a single simple form and provides a relative error less than 0.4% for $x \in (-\infty, +\infty)$, hence has been widely used in the device modeling community²⁸. The expression in Ref. 26 is given as

$$\begin{aligned} \mathcal{F}_{\frac{1}{2}}(x) &\approx \left(e^{-x} + \frac{3\sqrt{\pi}}{4} v^{-3/8} \right)^{-1}, \\ v &= x^4 + 50 + 33.6x(1 - 0.68 \exp[-0.17(x+1)^2]). \end{aligned} \quad (7)$$

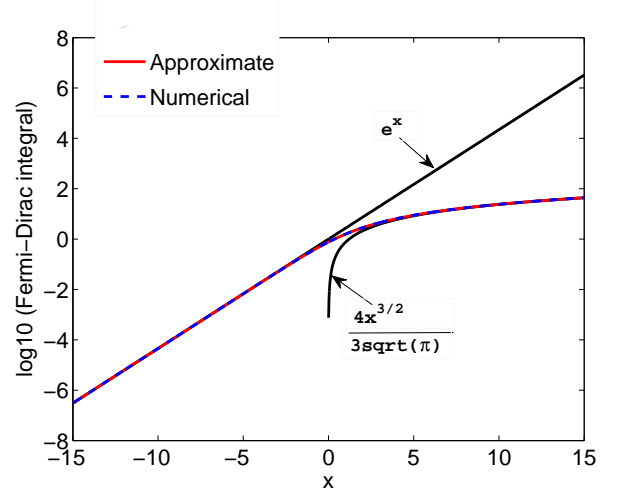


FIG. 2. (Color online) Fermi-Dirac integral of 1/2 order in logarithmic scale. Black curves are the asymptotic expansions, $\exp(x)$ for $x \rightarrow -\infty$, and $\frac{4x^{3/2}}{3\sqrt{\pi}}$ for $x \rightarrow +\infty$. The red solid curve labeled as Approximate is obtained using Eq. (7), while the blue dashed curve labeled as Numerical uses the numerical integration method in Ref. 21. The red solid and blue dashed curves do not show any visible difference, and agree well with the asymptotic forms for large $|x|$, whereas they differ significantly from the asymptotic forms in the region of $x \in (-4, 4)$.

The asymptotic expansion at $x \rightarrow -\infty$ leads to $\mathcal{F}_{1/2}(x) = \exp(x)$, implying that Eq. (3) becomes equivalent to Eq. (2), which is the case for non-degenerate semiconductors where $E_F \ll (E_C - [\text{a few } k_B T])$ for n , and $E_F \gg (E_V + [\text{a few } k_B T])$ for p . The asymptotic form at $x \rightarrow +\infty$ is $\mathcal{F}_{1/2}(x) = \frac{4x^{3/2}}{3\sqrt{\pi}}$, which corresponds to the strongly degenerate case near 0 K, where $E_F \gg (E_C + [\text{a few } k_B T])$ for n (i.e., the Fermi level is located within the conduction band), and $E_F \ll (E_V - [\text{a few } k_B T])$ for p (i.e., the Fermi level is inside the valence band). Figure 2 shows a comparison of the 1/2-order Fermi-Dirac integral calculated by different methods. It is clear that the approximate expression in Eq. (7) produces visually the same result as the numerical approach²¹, and follows the proper asymptotic forms for large $|x|$. In the small $|x|$ regime, neither of the asymptotic forms is valid. Since semiconductor DQD qubits are currently operated in this regime (corresponding to very low temperatures, mK to a couple of K), it is important to adopt a sufficiently accurate evaluation of the Fermi-Dirac integral. Due to the good accuracy and simplicity of Eq. (7), we implemented this form in QCAD for the FD statistics. In the actual implementation, we approximate $\mathcal{F}_{1/2}(x)$ by e^x for $x < -50$ to avoid numerical instability caused by the e^{-x} term in Eq. (7). Such large negative values of x can occur at very low temperatures, and this approximation results in no discernible loss of accuracy, as shown in Fig. 2.

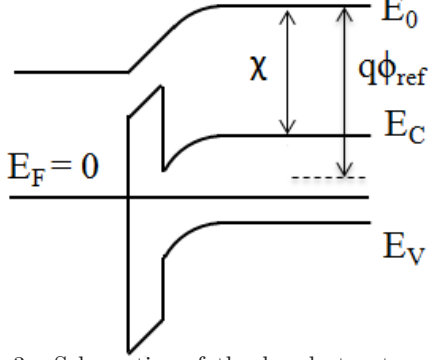


FIG. 3. Schematics of the band structure of a MOS-type device under zero bias.

2. The Reference Potential

Before solving the Poisson Eq. (1) for the electrostatic potential ϕ , one needs to relate ϕ to the band energies of the materials making up the device. One requirement for the electrostatic potential is that it must be continuous everywhere in a device. For a homo-junction device such as a PN silicon diode, E_C , E_V , and E_i (the intrinsic Fermi level) as functions of position are parallel to each other and continuous across the device, so it is natural to choose $-q\phi = E_i$, i.e., to solve for the inverse of intrinsic Fermi level. In an arbitrary hetero-junction structure, however, E_C , E_V , and E_i could all be discontinuous. Figure 3 shows a schematic of the band structure of a MOS-type device under zero bias illustrating the discontinuity of E_C and E_V . What is always continuous in arbitrary homo- and hetero-junction devices is the vacuum level indicated as E_0 in Fig. 3. Therefore, we choose ϕ to satisfy²⁹

$$-q(\phi - \phi_{ref}) = E_0 = E_C + \chi, \quad (8)$$

where ϕ_{ref} is a constant reference potential and χ is the electron affinity of a material. This choice implies that we are solving for the inverse of the vacuum level shifted by a constant value. While in theory different ϕ_{ref} values only change the resulting solution, ϕ , by a constant offset, in practice they can lead to different numerical convergence behavior during simulation. A good choice of ϕ_{ref} that has shown numerical robustness in devices containing silicon is to select as the reference potential the intrinsic Fermi level of silicon relative to the vacuum level, i.e., $q\phi_{ref} = E_0 - E_i(\text{Si})$. For a more detailed explanation of band diagrams schematics such as that in Fig. 3, we refer the reader to Ref. 30.

With the choice of ϕ in Eq. (8), we can rewrite E_C and E_V as

$$\begin{aligned} E_C &= -q(\phi - \phi_{ref}) - \chi, \\ E_V &= -q(\phi - \phi_{ref}) - \chi - E_g, \end{aligned} \quad (9)$$

with E_g being the band gap of a material. Then the n and p in Eq. (1) becomes a function of ϕ only, i.e., $n(\phi)$

and $p(\phi)$, as χ and E_g are material-dependent parameters and known for most semiconductors.

As QCAD does not solve the carrier transport (e.g., drift-diffusion) equations, all calculations must assume that thermal equilibrium (zero current flow) has been attained. The Fermi level, E_F , is taken to be a constant throughout any electrically-connected region of a device. The value of this constant is set by the voltages applied to the device. For example, if a voltage V_{sub} is applied to the substrate (right side) of the device in Fig. 3, E_F will become $-qV_{sub}$. The band structure shown in Fig. 3 corresponds to a MOS-type structure with metal gate (left side). If the structure instead had a highly-doped semiconductor gate (e.g., n^+ polysilicon gate) with applied voltage V_g , E_F in the gate region would be $-qV_g$ while E_F in the substrate region remains at $-qV_{sub}$.

3. Incomplete Ionization

When impurities are introduced into the semiconductor crystals, depending on the impurity energy level and the lattice temperature, not all dopants are necessarily ionized, especially at very low lattice temperatures (where DQD qubits are commonly operated). The ionized concentration for donors and acceptors is given by³¹

$$\begin{aligned} N_D^+ &= \frac{N_D}{1 + g_D \exp(\frac{E_F - E_D}{k_B T})}, \\ N_A^- &= \frac{N_A}{1 + g_A \exp(\frac{E_A - E_F}{k_B T})}, \end{aligned} \quad (10)$$

where E_D is the donor energy level, g_D is the donor ground state degeneracy factor, E_A is the acceptor energy level, and g_A is the acceptor ground state degeneracy factor. g_D is equal to 2 because a donor level can accept one electron with either spin or can have no electron when filled. g_A is equal to 4 because in most semiconductors each acceptor level can accept one hole of either spin and the impurity level is doubly degenerate as a result of the two degenerate valence bands (heavy hole and light hole bands) at the Γ point.

To write N_D^+ and N_A^- as a function of ϕ , we need to rewrite $E_F - E_D = E_F - E_C + E_C - E_D = E_F - E_C + E_d$ with E_d being the donor ionization energy, and $E_A - E_F = E_A - E_V + E_V - E_F = E_a + E_V - E_F$ with E_a being the acceptor ionization energy. The most common donors in bulk Si are phosphorus (P) and arsenic (As), which have ionization energies of $E_d = 46$ meV and 54 meV respectively³¹. The most common acceptor dopant in bulk Si is boron (B), which has $E_a = 44$ meV³¹.

Substituting Eq. (9) and $E_F - E_D$, $E_A - E_F$ into

Eq. (10), we get³²

$$\begin{aligned} N_D^+ &= \frac{N_D}{1 + g_D \exp\left(\frac{E_F + E_d - q\phi_{ref} + \chi + q\phi}{k_B T}\right)}, \\ N_A^- &= \frac{N_A}{1 + g_A \exp\left(\frac{-E_F + E_a + q\phi_{ref} - \chi - E_g - q\phi}{k_B T}\right)}. \end{aligned} \quad (11)$$

With these expressions, N_D^+ and N_A^- also become a function of ϕ , i.e., $N_D^+(\phi)$ and $N_A^-(\phi)$. Hence, the entire right hand side (RHS) of Eq. (1) can be written as a non-linear function of ϕ . Applying integration by parts and divergence theorem, we then rewrite the equation into the finite element (FE) weak form,

$$\begin{aligned} \int \epsilon_s \nabla \phi \cdot \nabla w d\Omega - \int_{\Gamma} \epsilon_s \nabla \phi \cdot \hat{\eta} w d\Gamma \\ - \int q[p(\phi) - n(\phi) + N_D^+(\phi) - N_A^-(\phi)] w d\Omega = 0, \end{aligned} \quad (12)$$

where w is the FE nodal basis function and the second term is a line integral over the simulation domain boundary with $\hat{\eta}$ being the unit normal vector of the surface element $d\Gamma$. The weak form is discretized using the Trilinos/Intrepid library, and the resulting discrete equation is solved by a nonlinear Newton solver also in Trilinos. Both the discretization library and the Newton solver were made directly available to QCAD through the Albany framework (cf. Fig. 1).

4. Boundary Conditions

An essential ingredient to the formulation of a differential equation are boundary conditions (BCs). QCAD supports three types of BCs: Dirichlet, Neumann, and Robin BCs. We will next discuss the implementation of these types in turn.

Dirichlet BCs are divided into two cases: (1) setting a voltage on the surface of a metallic region that borders insulator, and (2) setting a voltage on the surface of an Ohmic contact region which borders semiconductor. Case (1) is used for gate electrodes in field effect transistor (FET)-like devices, and the Dirichlet BC value ϕ_{ins} on the bordering insulator(s) is given by the simple expression

$$\phi_{ins} = V_g - \frac{\Phi_m - q\phi_{ref}}{q}, \quad (13)$$

with V_g being the applied gate voltage and Φ_m being the metal work function.

In the second case (used for Ohmic contacts in semiconductors), the potential on the bordering semiconductor surfaces is computed assuming thermal equilibrium and charge neutrality at the contacts. The calculation depends on carrier statistics and dopant ionization. For

MB statistics, the charge neutrality $n + N_A^- = p + N_D^+$ condition leads to

$$\begin{aligned} N_C \exp\left(\frac{E_F + q\phi - q\phi_{ref} + \chi}{k_B T}\right) + N_A^- \\ = N_V \exp\left(\frac{-E_F - q\phi + q\phi_{ref} - \chi - E_g}{k_B T}\right) + N_D^+. \end{aligned} \quad (14)$$

With complete ionization of dopants (i.e., $N_A^- = N_A$ and $N_D^+ = N_D$), the potentials at n-type and p-type Ohmic contacts are respectively given by

$$\begin{aligned} \phi_{ohm}^n &= \frac{q\phi_{ref} - \chi}{q} + \frac{k_B T}{q} \ln\left(\frac{N_D}{N_C}\right) + V_a, \\ \phi_{ohm}^p &= \frac{q\phi_{ref} - \chi - E_g}{q} - \frac{k_B T}{q} \ln\left(\frac{N_A}{N_V}\right) + V_a, \end{aligned} \quad (15)$$

where V_a is an externally applied voltage. When including incomplete ionization effect of dopants, we have for n-type and p-type semiconductors respectively,

$$\begin{aligned} N_C \exp\left(\frac{E_F - E_C}{k_B T}\right) &= \frac{N_D}{1 + 2 \exp\left(\frac{E_F - E_D}{k_B T}\right)}, \\ N_V \exp\left(\frac{E_V - E_F}{k_B T}\right) &= \frac{N_A}{1 + 4 \exp\left(\frac{E_A - E_F}{k_B T}\right)}. \end{aligned} \quad (16)$$

Let us denote $y_n = \exp\left(\frac{E_F - E_D}{k_B T}\right)$ and $y_p = \exp\left(\frac{E_A - E_F}{k_B T}\right)$. Then, by the definitions of E_a and E_d , we obtain the identities, $\exp\left(\frac{E_F - E_C}{k_B T}\right) = y_n \exp\left(\frac{-E_d}{k_B T}\right)$ and $\exp\left(\frac{E_V - E_F}{k_B T}\right) = y_p \exp\left(\frac{-E_a}{k_B T}\right)$. Substituting the identities into Eq. (16), we obtain

$$\begin{aligned} y_n &= -\frac{1}{4} + \frac{1}{4} \left[1 + \frac{8N_D}{N_C} \exp\left(\frac{E_d}{k_B T}\right) \right]^{1/2}, \\ y_p &= -\frac{1}{8} + \frac{1}{8} \left[1 + \frac{16N_A}{N_V} \exp\left(\frac{E_a}{k_B T}\right) \right]^{1/2}. \end{aligned} \quad (17)$$

From the definitions of y_n and y_p , the use of $-qV_a = E_F$, and Eq. (9), we can obtain the potentials that include dopant incomplete ionization effect at the n-type and p-type Ohmic contacts respectively as,

$$\begin{aligned} \phi_{ohm}^n &= \frac{-E_d + q\phi_{ref} - \chi}{q} + \frac{k_B T}{q} \ln(y_n) + V_a, \\ \phi_{ohm}^p &= \frac{-E_a + q\phi_{ref} - \chi - E_g}{q} - \frac{k_B T}{q} \ln(y_p) + V_a. \end{aligned} \quad (18)$$

At very low temperatures, the exponential terms in Eq. (17) could blow up numerically. To avoid numerical instability in QCAD, we approximate the $\ln(y_n)$ and $\ln(y_p)$ terms for very low temperatures as,

$$\begin{aligned} \ln(y_n) &= \frac{1}{2} \ln\left(\frac{N_D}{2N_C}\right) + \frac{E_d}{2k_B T}, \\ \ln(y_p) &= \frac{1}{2} \ln\left(\frac{N_A}{4N_V}\right) + \frac{E_a}{2k_B T}. \end{aligned} \quad (19)$$

For FD statistics and assuming complete ionization of dopants, we have

$$\begin{aligned} N_C \mathcal{F}_{\frac{1}{2}}\left(\frac{E_F - E_C}{k_B T}\right) &= N_D \quad \text{for n-type,} \\ N_V \mathcal{F}_{\frac{1}{2}}\left(\frac{E_V - E_F}{k_B T}\right) &= N_A \quad \text{for p-type.} \end{aligned} \quad (20)$$

To solve for the potentials at Ohmic contacts, in principle, we need to numerically solve Eq. (20) as the Fermi-Dirac integral does not have an analytic result. In QCAD, we use an approximate expression for the inverse of the $1/2$ order Fermi-Dirac integral, that is, given $u = \mathcal{F}_{1/2}(\eta)$, η is computed as³³,

$$\begin{aligned} \eta &= \frac{-\ln(u)}{u^2 - 1} + \frac{\nu}{1 + (0.24 + 1.08\nu)^{-2}}, \\ \nu &= \left(\frac{3\sqrt{\pi}u}{4}\right)^{2/3}. \end{aligned} \quad (21)$$

This approximation has an error of less than 0.6% for the entire η range. Using this expression, we obtain the BC potentials as,

$$\begin{aligned} \phi_{ohm}^n &= \frac{q\phi_{ref} - \chi}{q} + \frac{k_B T}{q}\eta + V_a, \\ \phi_{ohm}^p &= \frac{q\phi_{ref} - \chi - E_g}{q} - \frac{k_B T}{q}\eta + V_a, \end{aligned} \quad (22)$$

with η given in Eq. (21) where $u = N_D/N_C$ for n-type and $u = N_A/N_V$ for p-type. For the case of FD statistics and incomplete ionization, there exists no approximate analytic expressions for the BC potentials, and one has to solve a non-trivial nonlinear equation if want to be very accurate. In QCAD, we approximate this case using MD statistics with incomplete ionization and utilize the BC potentials given in Eq. (18).

Neumann BCs in finite element methods are used to specify how “flux” is conserved across boundaries. By default, all boundaries that are not given any other type of boundary conditions, assume implicit Neumann BCs which preserve the flux. In the case of the Poisson equation, the flux is $\epsilon_s \nabla \phi \cdot \hat{\eta}$, where $\hat{\eta}$ is the unit normal of the boundary surface. Thus, by default (i.e. when no other boundary condition is specified), $\epsilon_s \nabla \phi \cdot \hat{\eta} = 0$ on outer boundaries of the finite element mesh and $\epsilon_{s1} \nabla \phi_1 \cdot \hat{\eta}_1 = \epsilon_{s2} \nabla \phi_2 \cdot \hat{\eta}_2$ on internal boundaries. These two conditions are automatically satisfied in the finite element framework by setting the $\int \epsilon_s \nabla \phi \cdot \hat{\eta} w d\Gamma$ term to 0 in Eq. (12).

QCAD has the ability to specify non-flux-conserving Neumann BCs on specific boundaries such that the difference between the fluxes on either side of the boundary are equal to some specified constant value. Written mathematically, $(\epsilon_{s2} \nabla \phi_2 - \epsilon_{s1} \nabla \phi_1) \cdot \hat{\eta} = q\sigma_s$, where $\hat{\eta}$ is the unit normal vector of the interface pointing from material 2 to 1 and σ_s is the specified constant. Physically, σ_s is a surface charge density located at the boundary. Note that when $\sigma_s = 0$ (i.e. no surface charge), the boundary

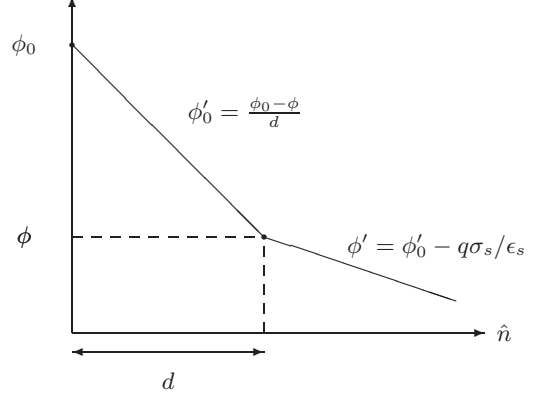


FIG. 4. The derivation of the Robin BC parameters used to set a value of the potential and a nearby surface charge on a single surface. This diagram shows a 1D cut along the direction \hat{n} normal to a surface element. ϕ_0 , σ_s , and d are given values, and ϕ is the variable being solved. The derivative ϕ' is along the direction normal to the surface. Combining the equations yields $\epsilon_s \phi' = \epsilon_s \frac{\phi_0 - \phi}{d} - q\sigma_s$, which takes the form of a Robin boundary condition.

condition reduces to the default flux-conserving condition. Within the finite element discretization in QCAD, this type of Neumann BC is implemented in the integral form

$$\int_{\Gamma_{cbc}} (\epsilon_{s2} \nabla \phi_2 - \epsilon_{s1} \nabla \phi_1) \cdot \hat{\eta} w d\Gamma_{cbc} = \int_{\Gamma_{cbc}} q\sigma_s w d\Gamma_{cbc}. \quad (23)$$

A major shortcoming of the Neumann BCs is their inability to characterize surface charge on (or extremely close to) an interface whose voltage is set by a Dirichlet BC. This is due to the simple fact that specifying both Dirichlet and Neumann BCs on the same surface overdetermines the problem. Yet, this is essentially what is needed to model a layer of charge that is stuck to one of the conducting gates (often polysilicon) used to control a device. One way around this technical difficulty is to place a layer of very thin finite-element cells around the charged gate and set a Neumann BC on the new surface lying a small distance away from the gate itself. This approach, however, suffers due to the thin finite elements adversely affecting convergence and their being hard to create in the first place. Instead, we use what are called Robin boundary conditions³⁴ to address the issue of charged gates. Robin BCs are similar to Neumann BCs but allow the flux at a surface to depend on the value of the solution (in this case the potential) there. Specifically, the Robin BC for an internal surface element can be written $(\epsilon_{s2} \nabla \phi_2 - \epsilon_{s1} \nabla \phi_1) \cdot \hat{\eta} = C + \alpha(\phi_2 - \phi_1)$, where C and α are fixed constants. At an external surface, we have $\epsilon_s \nabla \phi \cdot \hat{\eta} = C + \alpha\phi$. We would like to roll into a single boundary condition, a Dirichlet condition

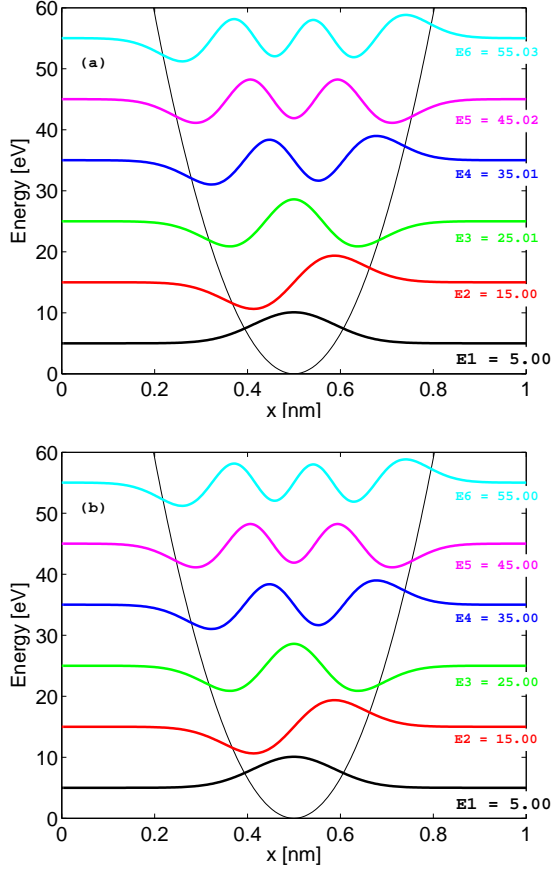


FIG. 5. (Color online) (a) Wave functions and energies obtained from the QCAD Schrodinger solver for a 1D parabolic potential well. (b) Analytic wave functions and energies for the same potential well. All the wave functions are scaled by the same factor for easy visualization. It is clear that QCAD wave functions and energies agree very well with the analytic results.

at one surface followed by a Neumann boundary condition at a parallel surface lying a very small distance away from the first surface. This can be done at an external surface using $C = \epsilon_s \phi_0 / d - q\sigma_s$ and $\alpha = \epsilon_s / d$, as shown in Fig. 4, which places a surface charge of $q\sigma_s$ a distance d away from a point at which ϕ is pinned to ϕ_0 . We somewhat arbitrarily choose $d = 10$ nm, which is much smaller than any of the mesh features (for semiclassical Poisson simulations) in our devices of interest. Robin BCs are enforced in an integral form, similar to that of Neumann BCs (cf. Eq. 23).

B. Schrodinger Solver

The time-independent single-particle effective mass Schrodinger equation takes the form of

$$-\frac{\hbar^2}{2} \nabla \cdot \left(\frac{1}{m^*} \nabla \psi(\mathbf{r}) \right) + V(\mathbf{r}) \psi(\mathbf{r}) = E \psi(\mathbf{r}). \quad (24)$$

The FE weak form of the equation is

$$\frac{\hbar^2}{2m^*} \left(\int \nabla \psi \cdot \nabla w d\Omega - \int_{\Gamma} \nabla \psi \cdot \hat{\eta} w d\Gamma \right) + \int V \psi w d\Omega - \int E \psi w d\Omega = 0 \quad (25)$$

The weak form is discretized by the FE method and the resulting eigenvalue problem $[H][\psi] = [E][\psi]$ is solved by the Trilinos eigensolver package called Anasazi¹⁶.

The Schrodinger solver supports two types of boundary conditions: Dirichlet and Neumann. For Dirichlet boundaries, $\psi = 0$. All other boundaries excluding Dirichlet are treated as Neumann BCs which, require $\frac{1}{m^*} \nabla \psi \cdot \hat{\eta} = 0$ on outer boundaries, and $\frac{1}{m^*} \nabla \psi \cdot \hat{\eta}$ being continuous (i.e., flux conservation) across material interfaces on internal boundaries. As in the Poisson solver, Neumann BCs are automatically satisfied in the FE framework by setting $\int_{\Gamma} \nabla \psi \cdot \hat{\eta} w d\Gamma = 0$ in Eq. (25). (The ability to set non-flux-conserving Neumann boundary conditions is absent in the Schrodinger solver since it would have no application in our work.)

Figure 5 shows a comparison between the QCAD Schrodinger solver and analytic results of the lowest six wave functions and energies for a 1D parabolic potential well. The QCAD and analytic results are in excellent agreement. The solver was also applied to 2D and 3D infinite potential wells. The obtained wave functions and energies were compared with the analytic results and excellent agreement was also observed.

One of the advantages of using the FE discretization over the finite difference discretization is that the continuities of ψ and $\frac{1}{m^*} \nabla \psi$ across heterojunctions are automatically satisfied in the former case, whereas they have to be explicitly enforced in the latter case. Specifically, when going from a homojunction device to a heterojunction device, the QCAD Schrodinger solver does not require any code change except setting the proper effective masses for the materials used. As an example, consider a 1D finite potential well that has a width of 20 nm, a potential height of 100 meV, and can have different effective masses for the well and barrier. Figure 6(a) shows the lowest three wave functions and energies obtained from QCAD for a homojunction device with the same effective mass for the well and barrier. In this case, the wave functions and their first derivatives are all continuous across the junctions. The wave functions and energies agree very well with the corresponding analytic results in Figure 2.15, Ref. 35. When the well and barrier have different effective masses, as shown in Fig. 6(b) for a heterojunction device, the wave functions are still continuous across the junctions, but their first derivatives are discontinuous due to the difference in effective masses.

C. Self-Consistent Schrodinger-Poisson Solver

In realistic quantum devices such as DQDs, we can divide the entire structure (relatively large) into semiclassi-

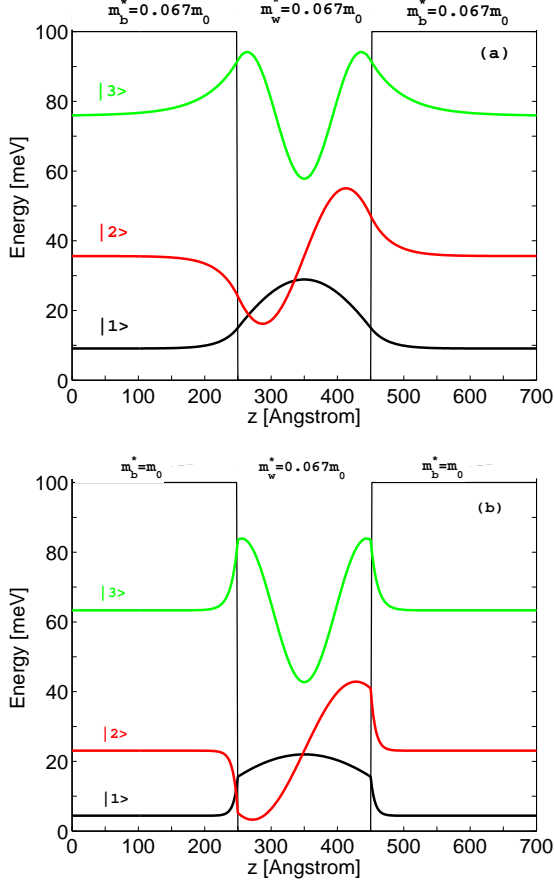


FIG. 6. (Color online) (a) Lowest three wave functions and energies obtained from the QCAD Schrodinger solver for the 1D finite potential well with $m_b^* = m_w^* = 0.067m_0$ where m_0 is the free electron mass. (b) Same as (a) except $m_b^* = m_0$ and $m_w^* = 0.067m_0$. All wave functions are scaled by the same factor for easy visualization.

cal and quantum regions. These regions are chosen such that in semiclassical regions, solving the nonlinear Poisson equation alone is often sufficient to obtain a good estimate of electrostatics, whereas in quantum regions, the Poisson and Schrodinger equations need to be coupled self-consistently for electrons (we focus on electrons only as they are used for qubit operation). The coupled two equations take the following form

$$-\nabla \cdot (\epsilon_s \nabla \phi) = q[p(\phi) - n(E_i, \psi_i) + N_D^+(\phi) - N_A^-(\phi)],$$

$$\frac{-\hbar^2}{2} \nabla \cdot \left(\frac{1}{m^*} \nabla \psi_i \right) + V(\phi, n) \psi_i = E_i \psi_i, \quad (26)$$

where the electron density n becomes a function of the i th energy level E_i and the envelope wave function ψ_i of the Schrodinger equation, while the potential energy V is a function of ϕ and n . The general expression for $n(E_i, \psi_i)$ is given by $\sum_i N_i |\psi_i|^2$, where the N_i term takes different expressions depending on confinement dimensionality.

1. Quantum Electron Density

In Si quantum devices, when we focus on those devices where the Si/other material (e.g., Si/SiO₂) interfaces are parallel to the [100] plane, the six equivalent conduction band minima of the bulk silicon are split into two groups due to the breaking of crystallographic symmetry, widely known as Δ_4 (fourfold degeneracy) and Δ_2 (double degeneracy) valleys, with Δ_2 valleys are lower in energy. At low temperatures, especially the operating temperatures for DQD qubits which are in the mK to a few Kelvin range, only the Δ_2 valleys are occupied by electrons; therefore, we consider the Δ_2 valleys only for Si devices in QCAD. Due to the ellipsoidal energy surfaces at the Δ_2 minima, the electron effective mass in the Schrodinger equation is different from that used in computing the electron density, and it depends on the confinement direction and the number of confined directions.

Two un-confined dimensions. In 1D-confined devices such as a 1D Si MOS capacitor, electrons are spatially confined in one direction (assumed x direction in QCAD) but free to move in the y and z directions. The Si/SiO₂ interface is in the $y-z$ plane and perpendicular to the longitudinal axis of the Δ_2 valleys. The coupled Schrodinger equation is 1D and given by

$$-\frac{\hbar^2}{2} \frac{d}{dx} \left(\frac{1}{m_l^*} \frac{d\psi_i(x)}{dx} \right) + V(\phi, n) \psi_i(x) = E_i \psi_i(x). \quad (27)$$

where m_l^* is the electron longitudinal effective mass of silicon. From Appendix A, the volume electron density n is computed as

$$n(E_i, \psi_i) = \sum_i N_i |\psi_i|^2 = \sum_i n_{2D,i} |\psi_i|^2$$

$$= \sum_i \left(2 \frac{m_t^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp \left(\frac{E_F - E_i}{k_B T} \right) \right] \right) |\psi_i|^2, \quad (28)$$

where m_t^* is the electron transverse effective mass of silicon and 2 accounts for the double degeneracy of the Δ_2 valleys. $|\psi_i(x)|^2$ is spatially normalized to 1, i.e., $\int |\psi_i(x)|^2 dx = 1$, and hence has the unit of 1/length. When $E_F > E_i$, the $\exp[(E_F - E_i)/(k_B T)]$ term in Eq. (28) can numerically go to infinity for very small T (mK to a few K), which can cause numerical instability. To avoid such problems, when the argument $(E_F - E_i)/(k_B T)$ is relatively large (e.g., > 100), we replace the $\ln[1 + \exp((E_F - E_i)/(k_B T))]$ term with $(E_F - E_i)/(k_B T)$.

One un-confined dimension. We next consider devices, such as quantum wire structures, where electrons are confined along two dimensions (assumed x and y directions in QCAD) and are free to move in the z direction (the wire direction). The Si/SiO₂ interface is perpendicular to the y axis and also the longitudinal axis of the

Δ_2 valleys. The coupled Schrodinger equation is 2D and given by

$$\begin{aligned} & -\frac{\hbar^2}{2} \frac{\partial}{\partial x} \left(\frac{1}{m_t^*} \frac{\partial \psi_i(x, y)}{\partial x} \right) - \frac{\hbar^2}{2} \frac{\partial}{\partial y} \left(\frac{1}{m_l^*} \frac{\partial \psi_i(x, y)}{\partial y} \right) \\ & + V(\phi, n) \psi_i(x, y) = E_i \psi_i(x, y). \end{aligned} \quad (29)$$

From Appendix A, the volume electron density n is computed as

$$\begin{aligned} n(E_i, \psi_i) &= \sum_i N_i |\psi_i|^2 = \sum_i n_{1D,i} |\psi_i|^2 \\ &= \sum_i \left[2 \left(\frac{2m_t^* k_B T}{\pi \hbar^2} \right)^{\frac{1}{2}} \mathcal{F}_{-\frac{1}{2}}(\eta_F) |\psi_i|^2 \right], \end{aligned} \quad (30)$$

where $\mathcal{F}_{-\frac{1}{2}}(\eta_F)$ is the Fermi-Dirac integral of -1/2 order. It is computed by using the approximate analytic expressions in Ref. 27, which have a very small error less than 0.001% in the entire η_F range. $|\psi_i(x, y)|^2$ is spatially normalized to 1, i.e., $\int |\psi_i(x, y)|^2 dx dy = 1$, and hence has the unit of $1/\text{length}^2$.

Zero un-confined dimensions. In devices such as quantum dot structures, electrons are spatially confined in all three directions and there are no (zero) dimensions in which they are free to move. The Si/SiO₂ interface is perpendicular to the z direction and also the longitudinal axis of the Δ_2 valleys. The coupled Schrodinger equation is 3D and given by

$$\begin{aligned} & -\frac{\hbar^2}{2} \frac{\partial}{\partial x} \left(\frac{1}{m_t^*} \frac{\partial \psi_i(x, y, z)}{\partial x} \right) - \frac{\hbar^2}{2} \frac{\partial}{\partial y} \left(\frac{1}{m_l^*} \frac{\partial \psi_i(x, y, z)}{\partial y} \right) \\ & - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \left(\frac{1}{m_l^*} \frac{\partial \psi_i(x, y, z)}{\partial z} \right) \\ & + V(\phi, n) \psi_i(x, y, z) = E_i \psi_i(x, y, z). \end{aligned} \quad (31)$$

The volume electron density n is computed as

$$n(E_i, \psi_i) = \sum_i N_i |\psi_i|^2 = \sum_i \left[\frac{4}{1 + \exp(\frac{E_i - E_F}{k_B T})} |\psi_i|^2 \right], \quad (32)$$

where 4 accounts for the double degeneracy of the Δ_2 valleys and that of the spin. $\psi_i(x, y, z)$ is normalized to 1 in the 3D quantum domain, and has the unit of $1/\text{length}^3$. When $E_i > E_F$, the $\exp[(E_i - E_F)/(k_B T)]$ term in Eq. (32) can blow up numerically. To avoid such problem, when $(E_i - E_F)/(k_B T)$ is relatively large (e.g., > 100), we replace the $[1 + \exp(\frac{E_i - E_F}{k_B T})]^{-1}$ term with $\exp(\frac{E_F - E_i}{k_B T})$.

All the above derivations are also applicable to other devices where the semiconductors have a single conduction band minimum located at the Γ valley such as GaAs-based devices, except that the valley degeneracy is 1 and a single electron effective mass is used in all the equations.

Next we discuss the potential energy term $V(\phi, n)$ in the coupled Schrodinger equation. It takes the form of

$$V(\phi, n) = q\phi_{ref} - \chi - q\phi + V_{xc}(n), \quad (33)$$

where $V_{xc}(n)$ is the exchange-correlation correction due to the Pauli exclusion principle in real many-electron systems. For the $V_{xc}(n)$ term, we use the well-known local density parameterization suggested by Hedin and Lundqvist³⁶ that has also been widely used by other authors^{28,37,38}. It is given as

$$\begin{aligned} V_{xc}(n) &= \frac{-q^2}{4\pi^2 \epsilon_s} [3\pi^2 n(\mathbf{r})]^{\frac{1}{3}} \left[1 + 0.7734x \ln \left(1 + \frac{1}{x} \right) \right], \\ x &= \frac{1}{21} \left(\frac{4\pi n(\mathbf{r}) b^3}{3} \right)^{-\frac{1}{3}}, \\ b &= \frac{4\pi \epsilon_s \hbar^2}{m_{xc}^* q^2}. \end{aligned} \quad (34)$$

Since this parameterization requires a scalar effective mass m_{xc}^* as input, we use an average mass for Si as suggested in Ref. 38,

$$\frac{1}{m_{xc}^*} = \frac{1}{3} \left(\frac{1}{m_l^*} + \frac{2}{m_t^*} \right). \quad (35)$$

2. Self-Consistency

The Schrodinger (S) and Poisson (P) equations in Eq. (26) have strong nonlinear coupling. They need to be solved self-consistently by certain iterative numerical schemes. Various iteration schemes³⁹⁻⁴⁶ have been proposed and used over the past few decades. Among them, three are notable: the under-relaxation method, the damped Newton method, and the predictor-corrector approach.

The under-relaxation scheme^{39,42} (also called convergence-factor or simple average method) solves the S and P equations in succession, and under-relaxes the electron density n or the electrostatic potential ϕ for the k th S-P outer iteration, using either a pre-set constant or an adaptively determined relaxation parameter $w^{(k)}$ (see the references for details). The advantage of this method is its simplicity. Its weakness is that the relaxation parameter $w^{(k)}$ is not known in advance and needs to be dynamically but heuristically readjusted during the course of iterations; if $w^{(k)}$ is too large, the iteration loop cannot reach convergence, whereas, if $w^{(k)}$ is too small, it takes too many iteration steps to achieve convergence.

The damped Newton method^{41,44} also solves the S and P equations in succession, but uses a damped Newton method⁴⁷ for the outer S-P iteration. Specifically, this approach starts from an initial guess $\phi^{(0)}$, solves the Schrodinger eigenvalue problem, computes quantum electron density $n^{(k)}$ according to section IIC 1, and then solves a linear P equation, obtained by linearizing the Poisson equation in Eq. (26) according to the Newton method⁴⁷ with an approximate Jacobian matrix; the potential $\phi_{out}^{(k)}$ from the linearized P equation is used to obtain $\phi_{in}^{(k+1)} = \phi_{in}^{(k)} + w^{(k)}(\phi_{out}^{(k)} - \phi_{in}^{(k)})$, which is then input to the S equation, and the procedure continues

until self-consistency is reached. Here, the $w^{(k)}$ damping parameter is not heuristic, but can be determined by the selection algorithm of the damped Newton method (cf. Ref. 47). k represents the k th Newton iteration and also the k th outer S-P iteration. The Jacobian matrix must be approximated because the quantum electron density n does not have explicit dependence on the potential ϕ . The approximate Jacobian matrix is obtained by simply assuming a semi-classical electron density expression in the P equation. The approach has shown reasonably robustness⁴¹, however, because of the approximate nature of the Jacobian, it often takes many Newton iterations (e.g., 50) to achieve sufficient self-consistent accuracy (e.g., ϕ is converged within 0.01 meV).

It is well-known that the under-relaxation³⁹, damped Newton⁴¹, and other similar iteration schemes^{40,43} do not necessarily lead to convergence, or take too many iterations to achieve it. These schemes have been used mostly in 1D S-P problems and in only a limited number of 2D applications, and one may rightly expect that they would have much more difficulty in achieving convergence in 3D S-P problems (e.g. in quantum dots). The key reason for the instability of these methods is that they do not physically address the strong nonlinear coupling between the S and P equations. In 1997 Trellakis et. al⁴⁵ proposed the predictor-corrector (p-c) iteration scheme based on a perturbation argument. Due to its solid physical groundings, the p-c method has shown fast and robust convergence behavior^{32,45,46}, and has been widely used in 2D and 3D simulations of various quantum semiconductor devices⁴⁸⁻⁵⁰. Given its excellent track record, we implemented this p-c method in QCAD for the self-consistent S-P loop.

The key feature of the p-c method is that it partially decouples the S and P equations by moving most nonlinearities into the nonlinear Poisson equation

$$-\nabla \cdot (\epsilon_s \nabla \phi) = q[p(\phi) - \tilde{n}(\phi) + N_D^+(\phi) - N_A^-(\phi)], \quad (36)$$

where $\tilde{n}(\phi)$ is an approximate expression for the quantum electron density $n(E_i, \psi_i)$, which has an explicit dependence on the potential ϕ (note the exact quantum density $n(E_i, \psi_i)$ does not have explicit dependence on ϕ). The nonlinear Poisson equation can be solved by a Newton method (the predictor step). The predicted result for \tilde{n} and ϕ from this equation is then corrected in an outer iteration step by the solution of Schrodinger equation (the corrector step).

The approximate quantum density \tilde{n} is obtained by using the first-order perturbation theory and the derivative property of Fermi-Dirac integrals⁴⁵. The resulting \tilde{n} expression is the same as the exact quantum density n given in Section II C 1, except that the argument in the Fermi-Dirac integral is modified to include an explicit dependence on ϕ . For 1D-confined Si devices, \tilde{n} is given

by

$$\tilde{n}(\phi) = \sum_i \left(2 \frac{m_t^* k_B T}{\pi \hbar^2} |\psi_i^{(k)}|^2 \times \ln \left[1 + \exp \left(\frac{E_F - E_i^{(k)} + q(\phi - \phi^{(k)})}{k_B T} \right) \right] \right), \quad (37)$$

where the superscripts (k) denote quantities obtained in the previous k th outer S-P iteration step (hence they are known quantities). For 2D-confined Si devices,

$$\tilde{n}(\phi) = \sum_i \left[2 \left(\frac{2m_t^* k_B T}{\pi \hbar^2} \right)^{\frac{1}{2}} |\psi_i^{(k)}|^2 \times \mathcal{F}_{-\frac{1}{2}} \left(\frac{E_F - E_i^{(k)} + q(\phi - \phi^{(k)})}{k_B T} \right) \right]. \quad (38)$$

For 3D-confined Si devices,

$$\tilde{n}(\phi) = \sum_i \left[\frac{4 |\psi_i^{(k)}|^2}{1 + \exp \left(\frac{E_i^{(k)} - E_F - q(\phi - \phi^{(k)})}{k_B T} \right)} \right]. \quad (39)$$

Note that there is a minus sign in the $q(\phi - \phi^{(k)})$ term in Eq. (39). In principle, once the self-consistent S-P loop is converged, this term should be numerically zero, which might suggest that the sign shall not matter. However, our experience with QCAD is that the minus sign is very important for the 3D-confined case to achieve self-consistent convergence; if we used a plus sign here, the outer S-P loop ran into numerical oscillations.

The self-consistent p-c procedure in QCAD is done in the following steps.

(1) Solve the semiclassical nonlinear Poisson equation, Eq. (1), using the Newton solver in Trilinos¹⁶, to obtain an initial potential $\phi^{(0)}$ and compute the initial total potential energy $V^{(0)}$ without the exchange-correlation correction V_{xc} .

(2) Solve the coupled Schrodinger equation for the k th ($k \geq 1$) S-P iteration step,

$$-\frac{\hbar^2}{2} \nabla \cdot \left(\frac{1}{m^*} \nabla \psi^{(k)} \right) + V^{(k-1)} \psi^{(k)} = E^{(k)} \psi^{(k)},$$

to obtain $E^{(k)}$ and $\psi^{(k)}$ (performed using an eigensolver available in Trilinos).

(3) Solve the coupled nonlinear Poisson equation with the approximate quantum electron density $\tilde{n}^{(k)}(\phi^{(k)}; \phi^{(k-1)}, E^{(k)}, \psi^{(k)})$,

$$-\nabla \cdot (\epsilon_s \nabla \phi^{(k)}) = q[p(\phi^{(k)}) - \tilde{n}^{(k)} + N_D^+(\phi^{(k)}) - N_A^-(\phi^{(k)})],$$

using the Trilinos Newton solver to obtain the updated potential $\phi^{(k)}$, and compute $\tilde{n}^{(k)}$ and $V^{(k)}$ including $V_{xc}(\tilde{n}^{(k)})$. Note we want to use the latest electron density to compute V_{xc} for good convergence.

(4) Check if $\|\phi^{(k)} - \phi^{(k-1)}\| < \delta$ for $k \geq 2$ everywhere in the device, with δ being a pre-defined tolerance often chosen as $1 \times 10^{-5} \text{V}$; if not, repeat steps (2) to (4).

It is clear from the above procedure that there is no under-relaxation step between two S-P iterations and the outer iteration reduces to a simple alternation between solving S and P equations. In addition, the Newton Jacobian matrix for the nonlinear P equation, Eq. (36), can be found analytically, avoiding the necessity of using an approximate Jacobian matrix in the damped Newton iteration scheme⁴¹. In terms of code implementation, the p-c method was very straightforward to implement in QCAD within the Albany framework. Here we emphasize that, because of the automatic differential capability in Trilinos, the Newton Jacobian matrix is computed directly by the code, and we do not need to derive the Jacobian matrix. More details on the implementation and the Albany code structure are found in Ref. 17.

3. Validation Example

To validate the self-consistent S-P solver, we performed simulations on two structures and compared with other simulation results. The first one is a 1D MOS Si capacitor with 4-nm oxide and $5 \times 10^{17} \text{ cm}^{-3}$ p-substrate doping. Figure 7 compares the Δ_2 -valley lowest four wave functions and energies in the capacitor obtained from QCAD and SCHRED¹⁵ at $T = 50 \text{ K}$, the lowest temperature allowed by SCHRED. (SCHRED is a 1D self-consistent Poisson-Schrodinger solver for MOS capacitors available on www.nanohub.org.) There are two simulation differences between QCAD and SCHRED: (i) QCAD applies the S-P solver to both Si and SiO₂ regions, leading to slight wave function penetration in the oxide ($x < 0$) as seen in Fig. 7a, while SCHRED assumes $\psi = 0$ at the Si/SiO₂ interface; (ii) QCAD considers the two Δ_2 valleys only, whereas SCHRED includes both the two Δ_2 and the four Δ_4 valleys. A typical effective mass of $0.5m_0$ is often assumed^{51,52} for SiO₂, where m_0 is the free electron mass. To minimize the wave function difference near the Si/SiO₂ interface due to the different boundary conditions imposed by the two tools, we used $0.005m_0$ as the SiO₂ effective mass for the QCAD simulations. The choice of setting $m_{ox}^* = 0.005m_0$ is because, at the Si/SiO₂ interface, QCAD applies the flux conservation condition of $\frac{1}{m_{ox}^*} \cdot \frac{d\psi}{dx}|_{ox} = \frac{1}{m_{si}^*} \cdot \frac{d\psi}{dx}|_{si}$, and in order to make ψ at the interface as close to 0 as possible (to be more consistent with SCHRED), we need small $\frac{d\psi}{dx}|_{ox}$, meaning small m_{ox}^* . At $T = 50 \text{ K}$, we expect that ignoring the higher energy Δ_4 valleys produces negligible effect on the Δ_2 -valley results, as only the Δ_2 -valley lowest subband is occupied by electrons at this low temperature. As expected, the wave functions and energies in Fig. 7 show excellent agreement between QCAD and SCHRED. The results also indicate that in this device, the exchange-correlation potential V_{xc} significantly increases the subband energy separation due to

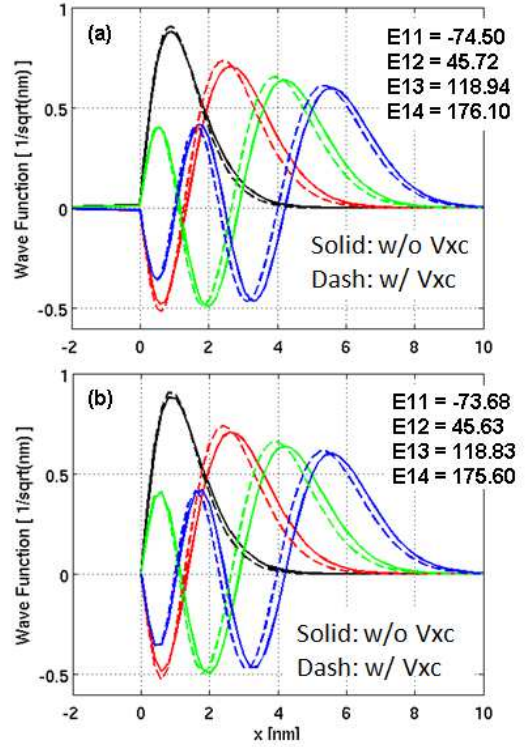


FIG. 7. (Color online) Δ_2 -valley lowest four subband wave functions and energies in a 1D MOS Si capacitor at $T = 50 \text{ K}$ and $V_g = 3 \text{ V}$ obtained from QCAD (a) and from SCHRED (b). The Si/SiO₂ interface is located at $x = 0$. The solid and dashed curves are obtained without and with the exchange-correlation effect, respectively. The subband energies in [meV], referenced from the Fermi level and including the V_{xc} effect, are denoted by E1*i*, where the “1” indicates the Δ_2 -valley and *i* indexes the subband (SCHRED’s labeling convention). For comparison, the corresponding energies without V_{xc} are -72.54, 26.71, 90.73, 144.69 for QCAD, and -71.76, 26.12, 89.22, 142.27 for SCHRED.

the many-body interaction (e.g., the separation between the lowest two subbands is increased from $26.71 + 72.54 = 99$ to $45.72 + 74.5 = 120 \text{ meV}$ when including V_{xc}), and it also somewhat compresses the wave functions as seen by the differences between the dash and solid curves in the figure, which agree with the observations in Ref. 28.

The second example is a gate-induced Si quantum wire structure from Ref. 53. Figure 8 shows the schematic diagram of the simulated 2D structure. For simulation purpose, the device is divided into quantum and semiclassical regions. The quantum regions include the 15-nm thick Si quantum and the 4-nm thick SiO₂ quantum regions denoted in the figure, where the self-consistent S-P solver is applied. The 15-nm and 4-nm were chosen such that the wave functions are essentially 0 at the boundaries of the quantum and non-quantum regions. The remaining Si and SiO₂ regions are treated as semiclassical, that is, only the Poisson equation with semiclassical carrier density is solved at each S-P iteration. The gate V_{g2} induces

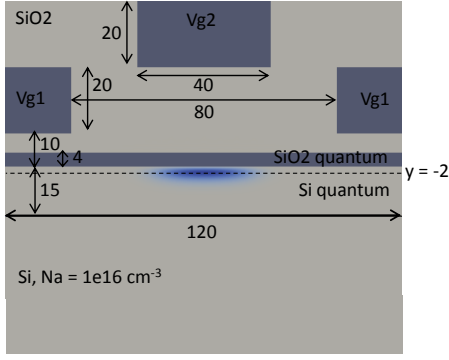


FIG. 8. (Color online) Schematic diagram of the simulated 2D structure with all dimensions given in nm. The blue 2D contour in the Si quantum region shows the Δ_2 -valley lowest subband wave function obtained from QCAD without the V_{xc} effect at $T = 10$ K, $V_{g1} = 0.8$ V, and $V_{g2} = 3.5$ V with all voltages referred to flat band. The dash line denotes the $y = -2$ nm location.

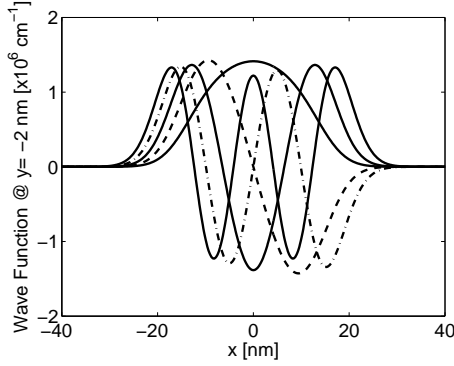


FIG. 9. (Color online) Δ_2 -valley lowest five subband wave functions along the $y = -2$ nm dash line in Fig. 8. These wave functions agree very well with Fig. 4(a) in Ref. 53.

electrons in the Si quantum region, while the V_{g1} gates are used to deplete electrons, hence an effective quantum wire is formed with the wire direction perpendicular to the 2D plane. The blue 2D contour in the Si quantum region shows the Δ_2 -valley lowest subband wave function obtained from QCAD without the effect of V_{xc} at $T = 10$ K, $V_{g1} = 0.8$ V, and $V_{g2} = 3.5$ V with all voltages referred to flat band. The peak of the wave function is located around the $y = -2$ nm dash line (the $y = 0$ location is at the Si-quantum/SiO₂-quantum interface). The lowest five subband wave functions along the $y = -2$ nm line are given in Fig. 9, which agree very well with Fig. 4(a) in Ref. 53. Given $T = 10$ K and $V_{g1} = 0.8$ V, we also performed QCAD S-P simulations for a range of V_{g2} voltages, integrated the electron density in the Si quantum region for each V_{g2} , and then plotted the subband energies as a function of integrated electron density to compare with the results in Ref. 53. Figure 10 compares the Δ_2 -valley lowest three subband energies as a function of integrated electron density in the Si quantum region

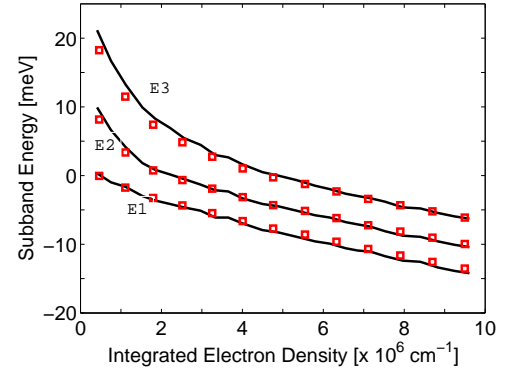


FIG. 10. (Color online) Δ_2 -valley lowest three subband energies as a function of integrated electron density in the Si quantum region. Black curves plot the data extracted from Fig. 3 in Ref. 53, while the red squares are the data obtained from QCAD. The energies are with respect to the Fermi level which is set to 0.

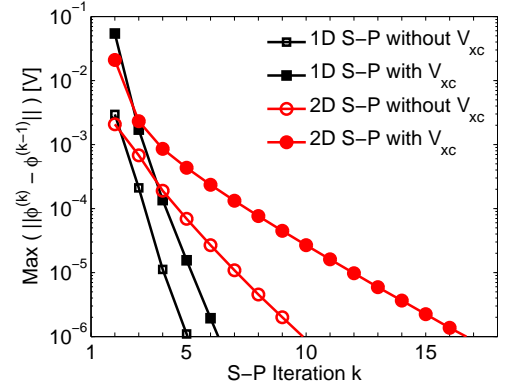


FIG. 11. (Color online) Convergence behavior of the self-consistent QCAD S-P solver for the 1D MOS Si capacitor and the 2D gate-induced Si quantum wire devices.

between QCAD and the reference. The agreement between them is very good considering the fact that the reference used a different S-P iteration scheme and did not mention if a fixed interface charge was used or not (no fixed charge was used in QCAD simulations).

For these 1D and 2D examples, their S-P convergence behavior from QCAD are plotted in Fig. 11, where the vertical axis is the maximum potential error between two outer iterations in the entire device including the semi-classical region (note that the corresponding convergence are not available from either the referenced tool or paper). We see that the predictor-corrector approach for the S-P iteration leads to fast and monotonic convergence in the two closed quantum systems; and the inclusion of the V_{xc} effect requires more iteration steps because of the stronger coupling that V_{xc} introduces between the P and S equations.

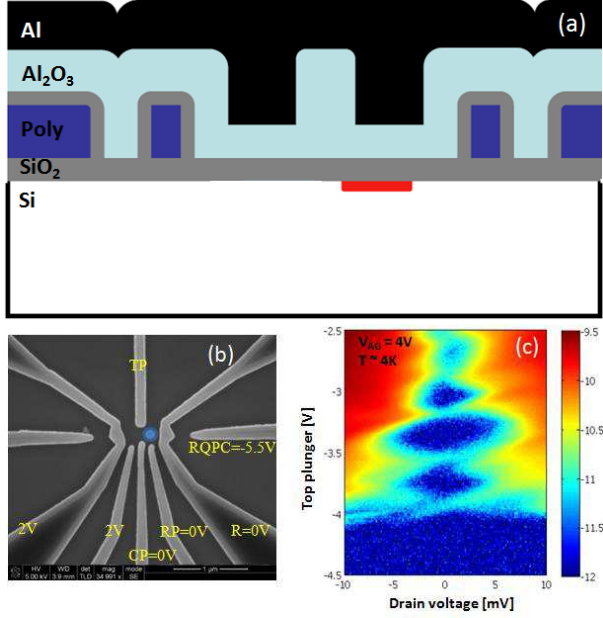


FIG. 12. (Color online) (a) Cross-section of an experimental Si DQD device, with the scanning electron micrograph (SEM) of the polysilicon depletion gates are shown in (b). The TP, CP, RP, R, and RQPC labels denote the depletion gates used to form the right dot. (c) Measured differential conductance for the right dot.

III. QD CAPACITANCE

In quantum dots, dot-to-gate capacitances are important quantities as they can be measured experimentally. Modeling of capacitances and comparison with experiment provide insight regarding the shape and location of a quantum dot and possible locations of defect charges. Capacitances can be computed in QCAD using either the semiclassical Poisson solver or the self-consistent S-P solver. In this section, we compare the capacitances found using both methods to those found experimentally.

Figure 12 shows the cross-section (a) and the SEM image of depletion gates (b) of an experimental Si DQD and the measured differential conductance for the right dot (c). The device is fabricated on a float-zone, high-resistivity (2-20 ohm-cm), p-type silicon substrate. The depletion gates are made of polysilicon degenerately doped with arsenic. A 60-nm Al_2O_3 insulation layer isolates a global Al top gate, which is positively biased to draw electrons into the quantum dot region. In the experiment, a quantum dot is formed on the right side of the nanostructure, schematically indicated as red square in (a) and as blue circle in (b). The 2-V bias applied to the left gates induces 2DEG below them and fills those regions with electrons such that confinement of electrons only occurs on the right side. Further details about the fabrication can be found elsewhere⁵⁴.

The quantum dot is held at the temperature of liquid helium, approximately 4 K, and a standard lock-

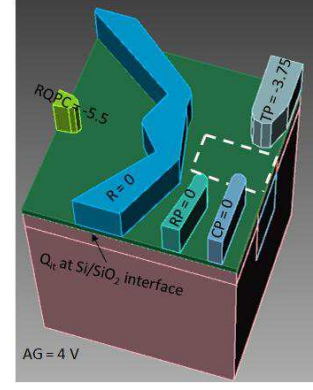


FIG. 13. (Color online) Single dot structure with polysilicon depletion gates, gate oxide, and substrate shown for computing dot-to-gate capacitances. The TP, CP, RP, R, and RQPC labels correspond to those in Fig. 12(b) and their experimentally applied voltages are given in unit of [V]. AG is the voltage applied to the global Al top gate that is not shown. Q_{it} is the fixed charge density at the Si/SiO₂ interface. The dashed white square denotes the top of the quantum dot region.

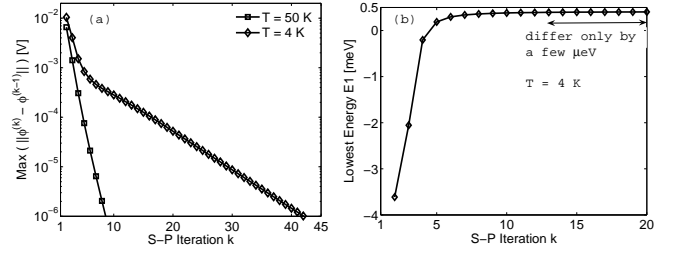


FIG. 14. (a) Convergence behavior of the maximum electrostatic potential error of the self-consistent 3D QCAD S-P solver applied to the Si quantum dot. (b) Convergence behavior of the ground state energy E_1 obtained by the S-P solver for the quantum dot at $T = 4$ K.

in technique is used to measure the differential conductance. The resulting conductance dependence on the top plunger gate bias as a function of the drain DC bias is shown in Fig. 12(c). The DC bias dependence of the Coulomb blockade through the quantum dot produces diamonds, typical of single dot behavior. The dot is observed to be empty after a last resonance around -4 V, indicative of the removal of the last electron. The total capacitance of the dot with only one electron is extracted from the charging energy, approximately 5 meV. Capacitances of the single-electron dot to other gates were extracted from the measured two-dimensional stability plots of the Coulomb blockade dependence on those gates in the device. The measured dot-to-gate capacitances are summarized in the experiment column of Table I.

To simulate capacitances of the right dot to the various gates and compare them with experiment, we transferred the depletion gates SEM image to a CAD drawing and created a single dot structure that replicates the right dot of the experimental DQD device. The created structure

is shown in Fig. 13 (for clarity, levels above the polysilicon depletion gates are not shown). The quantum region of this device is defined as a box volume of about 100-nm deep into the silicon substrate beneath the dashed white square. A positive voltage on the AG gate (not shown), which conformally covers the entire structure, induces a 2DEG in the silicon substrate under the Si/SiO₂ interface, while zero or negative voltages on the TP, CP, RP, and R gates deplete the 2DEG so that there is only one electron in the quantum dot. The RQPC gate is used for modulating a charge-sensing channel which can be used to measure the charge in the dot experimentally. Since the real device always contains an unknown number of defect charges in many locations, our QCAD simulations use an effective charge density Q_{it} at the Si/SiO₂ interface as a tuning parameter. Q_{it} is chosen such that the integrated electron density within the quantum region is equal to one under the given experimental voltages. Lumping all the possible effects of disorders into a single tuning parameter is clearly simplistic, but it avoids a time-consuming calibration process and gives a crude approximation of the real conditions while allowing for relatively fast calculation and analysis of the capacitances.

To compute the dot-to-gate capacitances including quantum effects, we apply the self-consistent 3D S-P solver to simulate the quantum region, while using the semiclassical Poisson solver for the rest of the device at each S-P iteration step. We see in Fig. 11 that the predictor-corrector scheme for the S-P loop leads to monotonic convergence for the 1D and 2D devices examined. For the 3D quantum dot, the S-P solver even with the V_{xc} effect also shows monotonic convergence as shown in Fig. 14(a) at lattice temperatures of 50 K and 4 K. Although the convergence of the potential is relatively slow at 4 K (42 iterations are needed to reach a maximum potential error of less than 1×10^{-6} V), the ground state energy E1 converges within much fewer iterations. As shown in Fig. 14(b), the energy E1 reaches an error of a few μ eV as early as the 13th iteration. In practice, we find that sufficient accuracy is obtained by running the simulation for 15 to 20 iterations at $T = 4$ K. For even lower temperatures such as in the milli-Kelvin range, more iteration steps may be required.

The calculated dot-to-gate capacitances are compared with experiment in Table I. The capacitance is computed using the static form of

$$C = \frac{dQ}{dV} = \frac{q(N_2 - N_1)}{dV},$$

$$N_{1,2} = \int \int \int n_{1,2}(x, y, z) dx dy dz,$$

where $N_{1,2}$ is the number of electrons in the quantum region for two given voltages with small difference dV (dV is fixed at 0.01 V) applied to the same gate. The QCAD Poisson simulations at 50 K were carried out to compare with the semiclassical results (not shown) obtained in the past by using the commercial Sentaurus Device tool¹⁴ which could not converge at temperatures below

TABLE I. Simulated and measured dot-to-gate capacitances for the single quantum dot shown in Fig. 13 operated in the one-electron regime.

	Experiment	QCAD Poisson	QCAD 3D S-P	QCAD Poisson	QCAD 3D S-P
T [K]	4	50	50	4	4
$Q_{it} [\times 10^{11} \text{ cm}^{-2}]$	unknown	-4.61	-4.51	-4.54	-4.43
e^- in the dot	1	0.96	0.96	0.95	0.95
dot-AG [aF]	2.37	3.98	4.33	4.38	4.98
dot-TP [aF]	0.48	0.33	0.37	0.37	0.44
dot-CP [aF]	0.54	0.86	0.96	0.95	1.12
dot-RP [aF]	0.29	0.64	0.72	0.71	0.84
dot-R [aF]	0.56	2.07	2.30	2.28	2.64

50 K. The semiclassical capacitances at 50 K obtained by Sentaurus are similar to those by the QCAD Poisson solver as expected. From the table, we see that, while the simulated capacitances do not match the measurement exactly, they have the correct order of magnitude, and the relative magnitude of values within one column correlates reasonably well with the other columns. We expect that the discrepancy between the simulated and measured values is due to the over-simplified treatment of defect charges in the simulations. In particular, the higher simulated dot-to-depletion-gate (CP, RP, and R) capacitances indicate that, there could be defect charges associated with these gates that need to be taken into account.

We observe that, for a given temperature and with a constraint of one electron in the dot, the capacitances calculated by the self-consistent 3D S-P solver are somewhat higher than the semiclassical values by the Poisson solver. One often supposes that quantum confinement will reduce capacitance, which is true if the number of electrons is allowed to change. In Table I, the electron number is fixed at approximately one in the dot. To understand the higher quantum capacitances (i.e., capacitances obtained by using the QCAD 3D S-P quantum solver), we performed analytical analysis on the 50-K case and also at 4 K. The electron density in the quantum dot region at $T = 50$ K is shown in Fig. 15. The conduction band 1D profiles along the depth up to 10-nm deep denoted by the white lines are plotted in Fig. 16. The choice of 10-nm depth is because electron density is small below the depth and can be neglected in the analysis. The 1D profiles are approximated by linear fits and used for analytical calculations. The profile obtained from the QCAD Poisson solver is well approximated by the linear fit, while the profile from the 3D S-P solver deviates from a linear fit near the Si/SiO₂ interface due to the inclusion of quantum confinement. The linear fits produce electric field of 7.045×10^6 V/m for the Poisson case and 8.256×10^6 V/m for the S-P case. Linear fits to the 1D band profiles allow for analytical calculations of semiclassical and quantum electron densities in triangular wells, which are

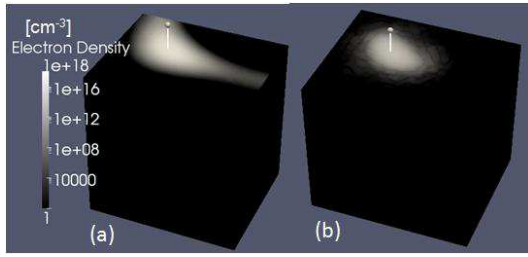


FIG. 15. Electron density in the quantum dot region of the device shown in Fig. 13 at $T = 50$ K obtained from the QCAD semiclassical Poisson solver (a), and from the self-consistent 3D S-P solver (b). The white vertical lines denote the locations of 1D cuts along the z direction.

described in Appendix B.

The analytical electron densities for the extracted 1D triangular wells are compared with the 1D cuts from the QCAD 3D solvers in Fig. 17. The agreement between the QCAD 1D cuts and the analytical solutions is exceptionally good, which indicates that the lateral (x - y plane) confinement is weak and the vertical (z direction) confinement dominates the solution in the Si quantum dot. The weak contribution of the lateral confinement can be seen from the ground state energy E_1 , which is equal to 9.79 meV (with respect to $E_F = 0$) from the QCAD 3D S-P solver, whereas it is 7.11 meV (the 25 meV shift is already included) from the analytic solution given in Appendix B. The slightly higher E_1 value from QCAD can be attributed to the lateral weak confinement. The very good agreement between QCAD and analytic indicates that we can use the analytical expressions to study how the capacitance varies between semiclassical and quantum calculations. From Appendix B, we know that the derivative of integrated electron density with respect to electric field, i.e., $d \int qn(z)dz/dF$ represents a capacitance per unit length. As shown in Fig. 18, the analytical quantum solution shows higher capacitance in the field range of interest, which is consistent with the observation in Table I. Intuitively, this is because given a fixed number of electrons in the dot, the quantum electron density is shifted away from the interface and much broader in space (see Fig. 17), which allows for somewhat higher charge variation for a given voltage change, leading to higher capacitance. Similar analysis and results also hold at $T = 4$ K.

IV. DQD OPTIMIZATION

The ability to tune a DQD device so that there are few electrons in its quantum dots while its tunnel barriers are controllable is very important for experimentally manipulating and measuring the device as a qubit. However, it is often challenging to achieve this, and the search for “good” DQD designs is difficult and time-consuming, due to the large space of possible depletion gate layouts and

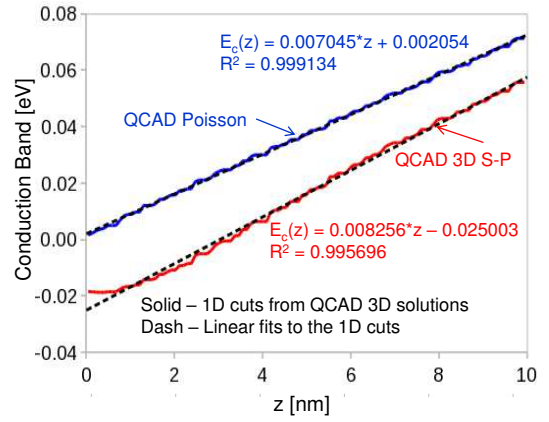


FIG. 16. (Color online) Conduction band profiles as a function of depth inside Si substrate starting from the Si/SiO₂ interface. The locations of the 1D cuts are indicated by the vertical lines in Fig. 15. Solid lines are the 1D cuts from QCAD 3D solutions, while dash lines are linear fits to the 1D cuts. The little wiggles on the solid lines are due to interpolation noise. The Fermi level E_F is set to 0.

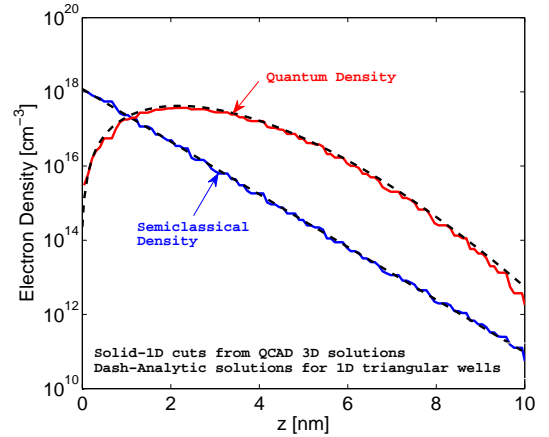


FIG. 17. (Color online) Electron density comparison between 1D cuts from the QCAD 3D solvers and the analytical solutions for the 1D triangular wells obtained in Fig. 16. The agreement between the QCAD 1D cuts and the analytical solutions is exceptionally good, which indicates that the lateral (x - y plane) confinement is weak and the vertical (z direction) confinement dominates the solution in the Si quantum dot.

the many choices related to which materials are used and the thickness of the various device layers. The ability to rapidly search over this large design space was one of the primary motivations for developing QCAD, and in this section we give an example of how the QCAD tool can be used to address the design of double quantum dot qubits.

Figure 19 shows the top view of four typical depletion gate patterns (transferred from SEM images) used in experimental DQD devices. For a given depletion gate layout, it often takes several months to find out experimentally if the device can work in the few-electron regime or not. Using QCAD’s optimization capabilities (pro-

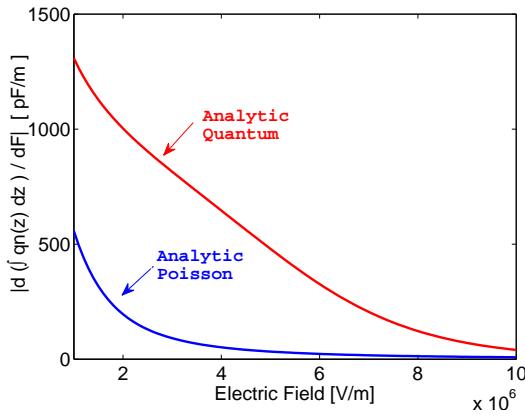


FIG. 18. (Color online) Analytical derivative of integrated electron density with respect to electric field in 1D triangular wells. The derivative has the unit of pF/m, representing a capacitance per unit length. The field range is selected such that it contains the fields extracted in Fig. 16 which keep the same number of electrons between Poisson and quantum calculations. For the entire field range, the analytical quantum solution shows higher capacitance, which is consistent with the observation in Table I.

vided by Dakota), we can perform optimization simulations of many device designs simultaneously on parallel clusters and can quickly determine whether a device is a few-electron DQD candidate or not. Each individual optimization would allow specified gate voltages to vary as adjustable parameters in order to meet one or multiple targets that are believed to be true of “good” DQD qubit designs. Such targets include (i) a given number of electrons in the left and/or right dot, e.g., one electron in the left dot in Fig. 19(a); (ii) specified values for the electron densities at multiple tunnel barriers, e.g., LTB, DB, and LQPCB in Fig. 19(a), which indicate whether the tunnel barriers can be turned on/off; (iii) the distance between a charge sensing constriction and a dot.

The locations of tunnel barriers in QCAD are dynamically detected through a saddle point searching algorithm based on the Nudged Elastic Band approach⁵⁵. This algorithm loosely solves the classical dynamics of a discretized elastic band stretched between two points. After the elastic band relaxes, the highest point on the band marks the saddle point. We have adopted several enhancements to the traditional nudged elastic band, including a climbing force added to the current highest point of the band (making it the “Climbing nudged elastic band” method) and a penalty term which discourages the band from making sharp turns weighted by an “anti-kinking” factor (a Gaussian smooth function). The initial and final points of the elastic band are determined either by direct user input or by maximizing a quantity of interest within a specified region. The latter case is used when we desire one end of the band to lie on a quantum dot whose location is a priori unknown: the starting (or ending) point is specified as the point of maximum

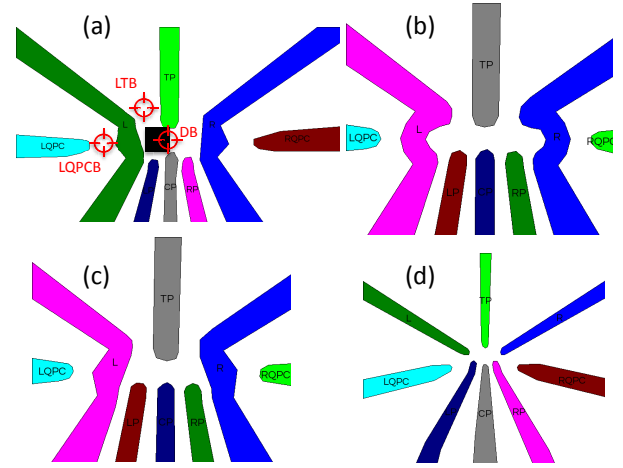


FIG. 19. (Color online) Top view of exemplary depletion gate patterns (transferred from SEM images) in experimental DQD devices. Each finger in the figures indicates a metal or polysilicon gate that can be set to a different voltage to form quantum dots. The red circles in (a) define the approximate locations for the left tunnel barrier (LTB), the dot barrier (DB), and the left QPC barrier (LQPCB). The black square in (a) denotes the left dot region.

electron density within the region where we expect the quantum dot to form.

Figure 20(a) shows a top view of the electron density at the Si/SiO₂ interface for the device shown in Fig. 19(a) after optimization. In this optimization, the depletion gates (TP, CP, LP tied to RP, L tied to R, LQPC tied to RQPC) and the top Al gate AG (not shown) voltages are allowed to vary, and the targets are to obtain one electron (integrated electron density close to one) in the left dot and simultaneously keep the electron densities in the LTB, DB, and LQPCB barriers at specified values (e.g., 10^{18} cm^{-3} at all three barriers). The gate voltages that meet these optimization targets are given in the caption of Fig. 20(a). For this device, the optimization results suggest that it can possibly operate in the one/few-electron regime with all tunnel barriers controllable. Experimental results have corroborated this observation, as measurements on this device have shown good few-electron characteristics. Optimization results for a different structure, shown in Fig. 20(b), suggest a different situation where the RQPCB barrier is completely shut off even when there are about 100 electrons in each dot. In this case, QCAD’s optimal solution could not meet its given targets simultaneously – one could find solutions for the target of having one electron in each dot, *or* for the target of setting electron densities to given values at the tunnel barriers, but not both, indicating that this device may not be able to reach the few-electron regime. This was also confirmed experimentally, as devices of this geometry did not display good few-electron behavior. These results show that the ability to perform device optimizations can help to efficiently narrow down

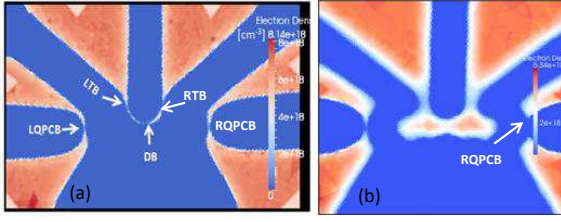


FIG. 20. (Color online) (a) View along the Si/SiO₂ interface of the resulting electron density after the optimizing the depletion gate voltages to obtain one electron in the left dot while keeping the electron densities in the LTB, DB, and LQPCB barriers at 10^{18} cm^{-3} . The device used is the same as that shown in Fig. 19(a) ($T = 0.2 \text{ K}$). The gate voltages that meet the optimization targets are $AG = 3.48 \text{ V}$, $TP = -0.74 \text{ V}$, $CP = -0.007 \text{ V}$, $LP = RP = -5.95 \text{ V}$, $L = R = -1.15 \text{ V}$, and $LQPC = RQPC = -2.41 \text{ V}$. (b) Top view of the electron density after optimization for a different DQD device, where there are about 100 electrons in each dot.

device candidates that meet essential requirements.

We have run optimizations such as the one described above on dozens of DQD device designs (some already fabricated, others not). Overall, these optimizations have helped us answer three DQD design questions that are critical in achieving few-electron QD behavior: (i) which devices allow one electron in each dot and also simultaneous control of tunnel barriers; (ii) for a given device, do tunnel barriers turn on before or after a dot has many electrons; (iii) what are the locations and shapes of the main dots and charge sensing constrictions.

V. CONCLUSION

We have developed a versatile finite-element-based tool called QCAD that is aimed at efficient simulation of multi-dimensional quantum devices, in particular, semiconductor multi-QD devices for use as qubits. QCAD is able to solve for the electrostatic potential with or without quantum effects (e.g. quantum confinement), and self-consistent single- and/or multi-electron wave functions and energies in 1D/2D/3D quantum devices. We have demonstrated robust convergence of the tool even at near-zero-Kelvin temperatures. The self-consistent quantum models in QCAD allow for analysis of quantum effects on device parameters of importance, and we have shown this for the case of dot-to-gate capacitances. Very high QCAD simulation throughput is achieved through pre- and post-processing scripting, distributed parallel computing capability and resources. QCAD simulations and optimizations of realistic multi-QDs allow for fast and valuable design comparison, optimization, and guidance to accelerate the experimental development of few-electron multi-QD qubits.

ACKNOWLEDGMENTS

This work was supported by the Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Appendix A

All derivations here assume single conduction band minimum and isotropic effective mass. In bulk devices (no spatial confinement), electrons are free to move in all three directions. Denote $\mathbf{r} = (x, y, z)$ and $\mathbf{k} = (k_x, k_y, k_z)$, and assume parabolic energy dispersion $E = E_C + \hbar^2 k^2 / (2m^*)$ with $k = |\mathbf{k}|$. The 3D electron density in bulk devices is given as

$$n_{3D} = \frac{2}{\Omega} \sum_{\mathbf{k}} f(E_F - E),$$

where 2 includes spin degeneracy and $f(E)$ is the Fermi-Dirac distribution, i.e., $f(E) = [1 + \exp(\frac{E - E_F}{k_B T})]^{-1}$. Making use of the sum-to-integral rule,

$$\frac{1}{\Omega} \sum_{\mathbf{k}} = \frac{1}{\Omega} \frac{\Omega}{(2\pi)^3} \int d^3 \mathbf{k},$$

and converting the integrand to spherical coordinates as the energy E only depends on the amplitude of \mathbf{k} ,

$$\frac{1}{\Omega} \sum_{\mathbf{k}} = \frac{1}{(2\pi)^3} \int d^3 \mathbf{k} = \frac{1}{(2\pi)^3} \int_0^{+\infty} 4\pi k^2 dk,$$

then the electron density becomes

$$n_{3D} = \frac{1}{\pi^2} \int_0^{+\infty} \frac{k^2 dk}{1 + \exp(\frac{E - E_F}{k_B T})}.$$

Since $E = E_C + \hbar^2 k^2 / (2m^*)$, we obtain

$$k^2 dk = \frac{1}{2} \left(\frac{2m^*}{\hbar^2} \right)^{\frac{3}{2}} \sqrt{E - E_C} dE,$$

then

$$n_{3D} = \frac{1}{2\pi^2} \int_{E_C}^{+\infty} \left(\frac{2m^*}{\hbar^2} \right)^{\frac{3}{2}} \frac{\sqrt{E - E_C}}{1 + \exp(\frac{E - E_F}{k_B T})} dE.$$

From the above expression, we can define the density of states (DOS) for bulk semiconductors, that is,

$$g_{3D}(E) = \frac{1}{2\pi^2} \left(\frac{2m^*}{\hbar^2} \right)^{\frac{3}{2}} \sqrt{E - E_C}.$$

With this definition, n_{3D} can be rewritten as

$$n_{3D} = \int_{E_C}^{+\infty} g_{3D}(E) f(E_F - E) dE,$$

which is the familiar expression widely found in solid state textbooks. Let $\varepsilon = (E - E_C)/(k_B T)$ and $\eta_F = (E_F - E_C)/(k_B T)$, then

$$\begin{aligned} n_{3D} &= \frac{1}{2\pi^2} \left(\frac{2m^* k_B T}{\hbar^2} \right)^{\frac{3}{2}} \int_0^{+\infty} \frac{\sqrt{\varepsilon} d\varepsilon}{1 + \exp(\varepsilon - \eta_F)}, \\ &= 2 \left(\frac{m^* k_B T}{2\pi \hbar^2} \right)^{\frac{3}{2}} \frac{2}{\sqrt{\pi}} \int_0^{+\infty} \frac{\sqrt{\varepsilon} d\varepsilon}{1 + \exp(\varepsilon - \eta_F)}, \\ &= N_C \mathcal{F}_{1/2}(\eta_F), \end{aligned} \quad (\text{A1})$$

where N_C is defined as the effective DOS in the conduction band and $\mathcal{F}_{1/2}(\eta_F)$ is the Fermi-Dirac integral of 1/2 order²².

In 1D-confined devices such as quantum well structures, electrons are free to move in two directions, whereas the third direction is spatially confined and forms quantized energy levels. Assume the free directions are along y and z , and the confined direction is along x . Denote $\mathbf{r} = (y, z)$ and $\mathbf{k} = (k_y, k_z)$. The parabolic energy dispersion becomes $E = E_i + \hbar^2 k^2/(2m^*)$, where E_i is the i th quantized energy level determined by the 1D Schrodinger equation in the confined direction x . The 2D electron density in the $y-z$ plane for a given E_i is defined as

$$n_{2D,i} = \frac{2}{A} \sum_{\mathbf{k}} f(E_F - E).$$

Following a procedure similar to the bulk case, we have

$$\begin{aligned} n_{2D,i} &= \frac{2}{A} \frac{A}{(2\pi)^2} \int d^2 \mathbf{k} f(E_F - E) \\ &= \frac{2}{(2\pi)^2} \int_0^{+\infty} 2\pi k dk f(E_F - E) \\ &= \frac{1}{\pi} \int_0^{+\infty} \frac{k dk}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} \\ &= \frac{m^*}{\pi \hbar^2} \int_{E_i}^{+\infty} \frac{dE}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} \\ &= \frac{m^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp\left(\frac{E_F - E_i}{k_B T}\right) \right] \\ &= \frac{m^* k_B T}{\pi \hbar^2} \mathcal{F}_0(\eta_F), \end{aligned} \quad (\text{A2})$$

where $\mathcal{F}_0(\eta_F)$ is the Fermi-Dirac integral of zero order with $\eta_F = (E_F - E_i)/(k_B T)$ here. In the final step of the derivation, the integral identity

$$\int \frac{dy}{1 + e^y} = \int \frac{e^{-y} dy}{1 + e^{-y}} = -\ln(1 + e^{-y})$$

is used. It is worthy of noting that one can obtain the 2D DOS from Eq. (A2), $g_{2D}(E) = m^*/(\pi \hbar^2)$, which is independent of energy.

In 2D-confined devices such as quantum wire structures, electrons are free to move in only one direction, whereas the other two directions are spatially confined and form quantized energy levels. Assume the free direction is along z , and the confined directions are along x and y . The parabolic energy dispersion becomes $E = E_i + \hbar^2 k_z^2/(2m^*)$, where E_i is the i th quantized energy level determined by the 2D Schrodinger equation in the confined directions. The 1D electron density in the z direction for a given E_i is defined as

$$n_{1D,i} = \frac{2}{L} \sum_{k_z} f(E_F - E).$$

Following a procedure similar to the bulk case, we have

$$\begin{aligned} n_{1D,i} &= \frac{2}{L} \frac{L}{2\pi} \int_{-\infty}^{+\infty} dk_z f(E_F - E) \\ &= \frac{2}{\pi} \int_0^{+\infty} dk_z f(E_F - E) \\ &= \frac{1}{\pi} \left(\frac{2m^*}{\hbar^2} \right)^{\frac{1}{2}} \int_{E_i}^{+\infty} \frac{(E - E_i)^{-\frac{1}{2}} dE}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)}. \end{aligned}$$

From this expression, we can define the 1D DOS as follows,

$$g_{1D}(E) = \frac{1}{\pi} \left(\frac{2m^*}{\hbar^2} \right)^{\frac{1}{2}} (E - E_i)^{-\frac{1}{2}}.$$

Let $\varepsilon = (E - E_i)/(k_B T)$ and $\eta_F = (E_F - E_i)/(k_B T)$, then

$$\begin{aligned} n_{1D,i} &= \frac{1}{\pi} \left(\frac{2m^* k_B T}{\hbar^2} \right)^{\frac{1}{2}} \int_0^{+\infty} \frac{\varepsilon^{-\frac{1}{2}} d\varepsilon}{1 + \exp(\varepsilon - \eta_F)} \\ &= \left(\frac{2m^* k_B T}{\pi \hbar^2} \right)^{\frac{1}{2}} \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \frac{\varepsilon^{-\frac{1}{2}} d\varepsilon}{1 + \exp(\varepsilon - \eta_F)} \\ &= \left(\frac{2m^* k_B T}{\pi \hbar^2} \right)^{\frac{1}{2}} \mathcal{F}_{-\frac{1}{2}}(\eta_F), \end{aligned} \quad (\text{A3})$$

where $\mathcal{F}_{-\frac{1}{2}}(\eta_F)$ is the Fermi-Dirac integral of -1/2 order²⁷.

Appendix B

Given the extracted 1D triangular well for the QCAD Poisson case in Fig. 16, $E_C(z) = 0.007045z + 0.002054$ eV = $qFz + b$ with the electric field $F = 7.045 \times 10^6$ V/m and the intercept $b = 0.002054$ eV, the semiclassical 3D electron density can be calculated as (assuming MB statistics for simplicity),

$$n(z) = N_C \exp\left(\frac{E_F - E_C}{k_B T}\right) = N_C \exp\left(\frac{E_F - qFz - b}{k_B T}\right),$$

with N_C given in Eq. (4). Integrating the electron density along z leads to

$$\int_0^{z_f} n(z) dz = \frac{N_C k_B T}{qF} \exp\left(\frac{E_F - b}{k_B T}\right) \left[\exp\left(\frac{-qF z_f}{k_B T}\right) - 1 \right],$$

which corresponds to a sheet electron density per unit area. z_f is an upper bound in z above which the electron density is negligible and it is set to 10 nm in correspondence with Fig. 16. Taking derivative of this expression with respect to the field gives

$$\begin{aligned} \frac{d \int_0^{z_f} qn(z) dz}{dF} &= qN_C \exp\left(\frac{E_F - b}{k_B T}\right) \\ &\times \left[\left(\frac{k_B T}{qF^2} + \frac{z_f}{F} \right) \exp\left(\frac{-qF z_f}{k_B T}\right) - \frac{k_B T}{qF^2} \right], \end{aligned} \quad (\text{B1})$$

which represents a capacitance per unit length.

Given the extracted 1D triangular well for the QCAD S-P case in Fig. 16, $E_C(z) = 0.008256z - 0.025003$ eV, we first shift the profile up to obtain $E_C(z) = 0.008256z$ eV = qFz with $F = 8.256 \times 10^6$ V/m, which leads to analytic solutions for the wave functions and eigen-energies, and then we shift the eigen-energies down by the same amount to obtain the actual energies. The i th analytic wave function for the 1D triangular well given by $E_C(z) = qFz$ takes the form of⁵⁶

$$\begin{aligned} \psi_i(z) &= A Ai[u(z)], \\ u(z) &= \left(\frac{2m_l^* qF}{\hbar^2} \right)^{\frac{1}{3}} \left(z - \frac{E_i}{qF} \right), \\ E_i &= \left(\frac{\hbar^2 q^2 F^2}{2m_l^*} \right)^{\frac{1}{3}} \left[\frac{3\pi}{2} \left(i - \frac{1}{4} \right) \right]^{\frac{2}{3}}, \\ A &= \left(\frac{2m_l^* qF}{\hbar^2} \right)^{\frac{1}{6}} \left[Ai'^2(\lambda_0) - \lambda_0 Ai^2(\lambda_0) \right]^{-\frac{1}{2}}, \\ \lambda_0 &= -\frac{E_i}{qF} \left(\frac{2m_l^* qF}{\hbar^2} \right)^{\frac{1}{3}}, i = 1, 2, \dots \end{aligned} \quad (\text{B2})$$

where A is the normalization factor, m_l^* is the longitudinal effective mass of Si, and $Ai(\lambda_0)$ is the Airy function. At $T = 50$ K and lower temperatures, and because of the low electron density in the DQD device, only the lowest subband is occupied, hence the volume quantum electron density can be computed as

$$n(z) = \frac{2m_t^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp\left(\frac{E_F - E_1}{k_B T}\right) \right] |\psi_1(z)|^2,$$

with 2 accounting for the Δ_2 -valley double degeneracy. E_1 here is equal to the E_1 in Eq. (B2) minus 0.025003 eV. As the wave function modulus is normalized, i.e., $\int_0^{+\infty} |\psi_1(z)|^2 dz = 1$, integrating the electron density leads to

$$\int_0^{+\infty} n(z) dz = \frac{2m_t^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp\left(\frac{E_F - E_1}{k_B T}\right) \right].$$

The derivative of the sheet electron density with respect to the field is given by

$$\begin{aligned} \frac{d \int_0^{+\infty} qn(z) dz}{dF} &= \frac{2m_t^* k_B T}{\pi \hbar^2} \left[1 + \exp\left(\frac{E_1 - E_F}{k_B T}\right) \right]^{-1} \\ &\times \frac{-1}{k_B T} \frac{dE_1}{dF}, \\ \frac{dE_1}{dF} &= \left(\frac{3\pi^2 \hbar^2 q^2}{16m_t^* F} \right)^{\frac{1}{3}}, \end{aligned} \quad (\text{B3})$$

which also represents a capacitance per unit length.

- ¹A. Morello, J. J. Pla, et al., *Nature* **467**, 687 (2010).
- ²M. G. Borselli, K. Eng, et al., *Appl. Phys. Lett.* **99**, 063109 (2011).
- ³W. H. Lim, C. H. Yang, F. A. Zwanenburg, and A. S. Dzurak, *Nanotechnology* **22**, 335704 (2011).
- ⁴G. Yamahata, T. Kodera, H. O. H. Churchill, K. Uchida, C. M. Marcus, and S. Oda, *Phys. Rev. B* **86**, 115322 (2012).
- ⁵L. A. Tracy, E. P. Nordberg, et al., *Appl. Phys. Lett.* **97**, 192110 (2010).
- ⁶T. M. Lu, N. C. Bishop, et al., *Appl. Phys. Lett.* **99**, 043101 (2011).
- ⁷E. A. Laird, J. M. Taylor, D. P. DiVincenzo, C. M. Marcus, M. P. Hanson, and A. C. Gossard, *Phys. Rev. B* **82**, 075403 (2010).
- ⁸T. S. Koh, J. K. Gamble, M. Friesen, M. A. Eriksson, and S. N. Coppersmith, *Phys. Rev. Lett.* **109**, 250503 (2012).
- ⁹M. Stopa, *Phys. Rev. B* **54**, 13767 (1996).
- ¹⁰S. N. Milicic, F. Badrieh, D. Vasileska, A. Gunther, and S. M. Goodnick, *Superlatt. and Microstruct.* **27**, 377 (2000).
- ¹¹M. Friesen, P. Rugheimer, D. E. Savage, M. G. Lagally, D. W. Weide, R. Joynt, and M. A. Eriksson, *Phys. Rev. B* **67**, 121301 (2003).
- ¹²D. V. Melnikov, J. Kim, L. X. Zhang, and J. P. Leburton, *IEEE Proc.-Circuits Devices Syst.* **152**, 377 (2005).
- ¹³B. Muralidharan, H. Ryu, Z. Huang, and G. Klimeck, *J. Comput. Electron.* **7**, 403 (2008).
- ¹⁴*Sentaurus Device User Guide*, Version D-2010.03, Synopsys Inc. and <http://www.quantumwise.com/>.
- ¹⁵<http://www.nanohub.org/>.
- ¹⁶<http://trilinos.sandia.gov>, <http://dakota.sandia.gov>, and <http://cubit.sandia.gov>.
- ¹⁷A. G. Salinger, R. P. Pawlowski, et al., *ACM Trans. Math. Software*, submitted (2013).
- ¹⁸X. Gao, E. Nielsen, R. P. Muller, R. W. Young, A. G. Salinger, N. C. Bishop, and M. S. Carroll, *Proc. of 15th IWCE*, DOI: 10.1109/IWCE.2012.6242832 (2012).
- ¹⁹E. Nielsen and R. P. Muller, *arXiv:1006.2735* (2010).
- ²⁰Christopher R. Anderson, *J. Comp. Phys.* **228**, 4745 (2009).
- ²¹N. Mohankumar and A. Natarajan, *Phys. Stat. Sol.(b)* **188**, 635 (1995).
- ²²J. S. Blakemore, *Solid-State Electron.* **25**, 1067 (1982).
- ²³See for example <http://mathworld.wolfram.com/Polylogarithm.html>
- ²⁴R. Kim and M. Lunstrom, *arXiv:0811.0116* (2008).
- ²⁵W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (3rd Ed., Cambridge University Press, 2007).
- ²⁶D. Bednarczyk and J. Bednarczyk, *Phys. Lett. A* **64**, 409 (1979).
- ²⁷P. Van Halen and D. L. Pulfrey, *J. Appl. Phys.* **57**, 5271 (1985); *ibid*, *J. Appl. Phys.* **59**, 2264 (1986).
- ²⁸D. Vasileska, *Solving the Effective Mass Schrodinger Equation in State-of-the-Art Devices*, <http://nanohub.org>.
- ²⁹M. S. Lundstrom, R. J. Schwartz, and J. L. Gray, *Solid-Stat. Electron.* **24**, 195 (1981).
- ³⁰IH. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, *J. Appl. Phys.* **68**, 4071 (1990).
- ³¹S. M. Sze and Kwok K. Ng, *Physics of Semiconductor Devices* (3rd Ed., John Wiley and Sons Ltd., 2007).

- ³²A. Trellakis and U. Ravaioli, *Comput. Methods Appl. Mech. Engrg.* **181**, 437 (2000).
- ³³N. G. Nilsson, *Phys. Stat. Solidi.(a)* **19**, K75 (1973).
- ³⁴Ronald B. Guenther and John W. Lee, *Partial Differential Equations of Mathematical Physics and Integral Equations* (Dover Publications Inc., 1996).
- ³⁵P. Harrison, *Quantum Wells, Wires and Dots* (2nd Ed., John Wiley and Sons Ltd., 2005).
- ³⁶L. Hedin and B. I. Lundqvist, *J. Phys. C: Solid State Phys.* **4**, 2064 (1971).
- ³⁷F. Stern and S. Das Sarma, *Phys. Rev. B* **30**, 840 (1984).
- ³⁸A. Trellakis and U. Ravaioli, *Solid State Electronics* **48**, 367 (2004).
- ³⁹F. Stern, *J. Comput. Phys.* **6**, 56 (1970).
- ⁴⁰C. Moglestue, *J. Appl. Phys.* **59**, 3175 (1986).
- ⁴¹S. E. Laux and A. C. Warren, *IEDM Technical Digest*, 567 (1986); S. E. Laux and F. Stern, *Appl. Phys. Lett.* **49**, 91 (1986).
- ⁴²T. Kerkhoven, A. G. Galick, U. Ravaioli, J. H. Arends, and Y. Saad, *J. Appl. Phys.* **68**, 3461 (1990).
- ⁴³J. Sune, P. Olivo, and B. Ricco, *J. Appl. Phys.* **70**, 337 (1991).
- ⁴⁴J. H. Luscombem, A. M. Bouchard, and M. Luban, *Phys. Rev. B* **46**, 10262 (1992).
- ⁴⁵A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli, *J. Appl. Phys.* **81**, 7880 (1997).
- ⁴⁶A. Trellakis, T. Andlauer, and P. Vogl, *Large Scale Scientific Computing* (Springer-Verlag Berlin Heidelberg, 2006).
- ⁴⁷R. E. Bank and D. J. Rose, *SIAM J. Numer. Anal.* **17**, 806 (1980).
- ⁴⁸G. Curatola, G. Iannaccone, *Comput. Mater. Sci.* **28**, 342 (2003).
- ⁴⁹H. R. Khan, D. Mamaluy, and D. Vasileska, *IEEE Trans. Electron. Devices* **54**, 784 (2007).
- ⁵⁰H. Wang, G. Wang, S. Chang, and Q. Huang, *Micro & Nano Lett.* **4**, 122 (2009).
- ⁵¹P. M. Schneider and W. B. Fowler, *Phys. Rev. Lett.* **36**, 425 (1976).
- ⁵²Y. P. Li and W. Y. Ching, *Phys. Rev. B* **31**, 2172 (1985).
- ⁵³S. E. Laux and F. Stern, *Appl. Phys. Lett.* **49**, 91 (1986).
- ⁵⁴E. P. Nordberg, G. A. Ten Eyck, et al., *Phys. Rev. B* **80**, 115331 (2009).
- ⁵⁵G. Henkelman, B. P. Uberuaga, and H. Jonsson, *J. Chem. Phys.* **113**, 9901 (2000).
- ⁵⁶<http://www.iue.tuwien.ac.at/phd/gehring/node116.html>, and http://ecee.colorado.edu/~bart/book/book/chapter1/ch1_2.htm.